DOT/FAA/CT-92/12, III

FAA Technical Center
Atlantic City International Airport
N.J. 08405

Reliability Assessment at Airline Inspection Facilities, Volume III: Results of an Eddy Current Inspection Reliability Experiment

May 1995

Final Report

This document is available to the public
through the National Technical Information Service,
Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/CT-92/12, III	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle RELIABILITY ASSESSMENT AT AIRLINE INSPECTION FACILITIES VOL. III RESULTS OF AN EDDY CURRENT INSPECTION RELIABILITY EXPERIMENT		Report Date May, 1995
		6. Performing Organization Code
7. Authors Floyd Spencer, Donald Schurman		8. Performing Organization Report No.
9. Performing Organization Name and Address Sandia National Laboratories Aging Aircraft NDI Validation Center Albuquerque, New Mexico		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTFA03-91-A-00018
12. Sponsoring Agency Name and Address US Department of Transportation Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code AAR 430

15. Supplementary Notes

Dr. Christopher Smith, FAA Technical Monitor
Dr. Patrick Walter, SNL, Aging Aircraft Project Manager — Volume III of III

16. Abstract

The Aging Aircraft NDI Development and Demonstration Center (AANC) at Sandia National Laboratories is charged by the FAA to support technology transfer, technology assessment, and technology validation. A key task facing the center is to establish a consistent and systematic methodology to assess the reliability of inspections through field experiments. This task is divided into three major areas: reliability of eddy current lap splice inspections at transport aircraft maintenance facilities, reliability of inspection at commuter aircraft maintenance facilities, and reliability of inspection associated with visual inspection of aircraft structural parts.

Volume III is the third in a series of three describing the planning, execution, and results of an eddy current inspection field experiment. The experiment was taken to nine facilities, and five inspections were performed at each facility. All inspections took place in the environment in which actual aircraft were inspected, were accomplished using the same equipment that would be used in actual aircraft inspections, and were performed by the same people who would do actual aircraft inspections. This document provides a detailed description of the results of that eddy current inspection reliability experiment (ECIRE).

The performance results of the inspections are summarized in the form of probability of detection (PoD) curves and are compared to a baseline established in a laboratory environment. From the laboratory inspections conducted as a baseline, the 90th percentile point on the PoD for inspections on a bare surface is estimated to be in the 0.06- to 0.07-inch range. As a comparative, the same percentile average from the field data was nominally 0.09 inches.

17. Key Words nondestructive inspection eddy current inspection reliability	18. Distribution Statement Document is avialable to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	175	

Form DOT F1700.7 (8-72) — Reproduction of completed page authorized

PREFACE

In August 1991, a major center with emphasis on validation of nondestructive inspection (NDI) techniques for aging aircraft was established at Sandia National Laboratories (SNL) by the Federal Aviation Administration (FAA). This center is the Aging Aircraft NDI Validation Center (AANC). It resides at the Albuquerque International Airport in a hangar leased from the City of Albuquerque, New Mexico. The FAA Interagency Agreement, which established this center, provided the following summary tasking statement: "The task assignments will call for Sandia to support technology transfer, technology assessment, technology validation, data correlation, and automation adaptation as ongoing processes."

As one of its first projects AANC established a working group consisting of personnel from Sandia National Laboratories, Science Application International Corporation (SAIC), and AEA Technology. The working group was charged with designing and implementing an experiment to quantify the reliability associated with detecting a crack originating within fastener holes in thin aluminum structure using high-frequency eddy current inspection methods.

The result of AANC's efforts is a three volume document, Reliability Assessment at Airline Inspection Facilities, which details an experimental concept for inspection reliability assessment and a specific experiment designed to determine probability of detection (PoD) curves associated with eddy current inspections. The protocol for the experiment was developed first as a generic protocol then as a specific eddy current lap splice inspection protocol. The generic protocol is presented in Volume I: A Generic Protocol for Inspection Reliability Experiments, and the specific eddy current experiment protocol is presented in Volume II: Protocol for an Eddy Current Inspection Reliability Experiment. The results and analysis of the experiment are presented here in the third volume, Volume III: Results of an Eddy Current Inspection Reliability Experiment.

ACKNOWLEDGMENTS

Many people have been instrumental in making this program possible. Those involved in the planning and review of the program prior to conducting the field experiment were acknowledged in Volume II. The efforts of that group would have come to naught, however, without the cooperation of the aviation NDI community. Specifically, the NDI foremen and the inspectors within several facilities gave their time and efforts to provide meaningful data. Those facilities are American Airlines in Tulsa, Oklahoma; Dalfort Aviation in Dallas, Texas; Aloha Airlines in Honolulu, Hawaii; Tramco in Evergreen, Washington; Alaska Airlines in Seattle, Washington; United Airlines in San Francisco, California; Delta Airlines in Atlanta, Georgia; US Air in Winston-Salem, North Carolina; and Miami NDT in Opa-Locka, Florida.

The experimental protocols were provided by AEA Technology. Ron Smith and Bob Murgatroyd of AEA, as well as Chris Smith, FAA Technical Center, were extremely helpful in exercising and critiquing these protocols before the experiment traveled to the facilities.

Science Applications International Corporation (SAIC) provided the logistic support to move the experiment around the country. Mike Ashbaugh of SAIC stayed on top of the logistics and assured that the hardware and the personnel came together at the various facilities on schedule. Jay Bustamante assisted Mike in the logistics task. Tim MacInnis of SAIC was instrumental in characterizing the test specimens. He also monitored the inspections at the facilities along with Don Schurman. Marty Taylor visited several of the facilities during the progress of the experiment and provided helpful insights.

Sandia National Laboratories was responsible for the overall program design and implementation as well as analysis of experimental results. Dennis Roach was responsible for the fabrication of the test specimens and Mike Nusser designed the test specimen frame supports. Pat Walter, as manager of the Aging Aircraft Program at Sandia, assisted throughout the project.

Special thanks are due to several people for reviewing an initial draft of this report. Specifically, Steve LaRiviere and Mike Hutchinson of Boeing Commercial Aircraft Group, Ward Rummel of Martin Marietta, and Chris Smith of the FAA Technical Center all provided extensive feedback. Their efforts have helped to improve the report.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

An experiment was designed to assess the reliability of detecting a crack originating within fastener holes in thin aluminum structure using high-frequency eddy current inspection methods. The test samples used were of the same construction as used on Boeing 737 lap splices. The test specimens had well-characterized crack lengths. The experiment was taken to nine facilities. Five inspections were performed at each facility. All inspections took place in the environment in which actual aircraft are inspected, were accomplished using the same equipment that would be used in actual aircraft inspections, and were performed by the same people who would do actual aircraft inspections.

The performance results of the inspections are summarized in the form of Probability of Detection (PoD) curves and are compared to a baseline established in a laboratory environment. From the laboratory inspections conducted as a baseline, the 90th percentile point on the PoD for inspections on a bare surface is estimated to be in the 0.06 to 0.07 inch range. As a comparative, the same percentile average from the field data was nominally 0.09 inch.

There was substantial inspector to inspector variation in performance. Some inspectors were capable of producing results similar to the laboratory results even in the presence of the operating environment distractions. The 90th percentile probability of detection point for the poorest performing inspectors was in the 0.14 to 0.18 inch range. False call rates were low for most of the inspectors. However, three inspectors exceeded 8 percent in false call rates with 90 percent probability of detection being achieved at 0.107, 0.087, and 0.146 inch in these three cases.

Facility differences were significant. Implementation of procedures and instrument-specific training were two factors identified as being highly variable between facilities.

Other factors that affect inspection performance include surface condition at time of inspection (painted or bare), crack orientation, and "accessibility." The surface condition accounted for shifts of approximately 0.01 inch in the PoD curves, with the painted surface having a lower probability of detection. The off-angle cracks of 11 and 22 degrees had lower probability of detection than did the horizontal cracks. The more inaccessible the lap splice, the lower the probability of detection; however, the difference was slight.

An overall background miss rate of approximately 0.024 was observed that was independent of crack lengths. This background miss rate changed with the condition of the inspection surface, with the painted panels having an increased level of crack misses that were independent of the crack length.

It was shown that inspectors operating in the maintenance environment, using Boeing procedures and a variety of equipment, are capable of routinely finding cracks as small as 0.100 inch. For most of the inspections (~62 percent or 28 of 45) the estimated detection rate for 0.100 inch cracks exceeded 0.90. A minority of the inspections (~18 percent or 8 of 45) had detection rates lower than 0.80 for cracks with lengths of 0.100 inch.

1. INTRODUCTION.

This document is the third in a three volume series addressing the quantification of inspection reliability at aircraft inspection facilities [1,2]. The results of conducting the experimental program proposed in the second volume [2] are presented and discussed.

1.1 PROGRAM PURPOSE.

Inspection of aircraft, especially aging aircraft, is a matter of national concern. Aircraft manufacturers, airlines, and the Federal Aviation Administration (FAA) have all established programs to address this concern.

High-frequency eddy current inspections are an integral part of routine maintenance checks and directed checks for surface fatigue cracks in airplane skins. The objective of the Eddy Current Inspection Reliability Experiment (ECIRE) proposed in volume II was to evaluate the capability and reliability of eddy current inspection procedures as they are done routinely at airline maintenance and inspection facilities [2]. Besides data needed to quantify the capability and reliability associated with the inspection process, data were also gathered on the facility environment and on the training and background of inspectors. Trained monitors traveled with the experiment and recorded not only the inspection results, but also observations on the maintenance environment and procedures. The intent was not only to be able to quantify the capability and reliability associated with the inspection process, but also to provide some insight into how the process could potentially be improved.

1.2 PRESENTATION FORMAT.

In section 2 the overall ECIR Experiment plan is reviewed. Specific information not available at the time of the planning presented in volume II [2], such as flaw sizes in the test specimens and facilities visited, is also presented in section 2. Observed characteristics of the facilities and the inspectors participating in the experiment are discussed in section 3. Section 4 contains an analysis of detection data gathered in the laboratory and in the field. Probability of detection fits to the field data are presented, as are Relative Operating Characteristic (ROC) curves. We discuss and summarize the results in section 5.

The reader who is interested in background and general results, without the detail, can obtain it by reading sections 2.2, 2.3, 3, and 5.

1.3 USE AND LIMITATIONS OF DATA.

The inspection data behind the analysis presented in this report consists primarily of crack lengths and whether a given inspector marked the rivet site containing that crack as flawed. Experimental design factors, such as the condition of the inspection surface (bare or painted) and the locations of flaws during the inspection task, are also a part of the database. The type of information gathered enable probability of detection curves to be fit to individual inspections, as well as a direct assessment of factors included in the experimental design.

A request was made through the Federal Aviation Administration Technical Center and the Air Transport Association of America for participants in this research. There was no compulsion to participate and the participants can not be considered a random sample of maintenance and inspection facilities. Similarly, the inspectors within a facility were designated by supervisors within the facility and can not be considered a random sample. It should be noted, however, that for a majority of the facilities, most, if not all inspectors participated.

Facility and inspector characteristics were noted during the experimentation. The observations constitute a part of the database. However, the nature of facility and individual inspector characteristics were not controlled in the experiment. As is usual with this type of data, influencing factors can be hypothesized, but statistical significance is not likely and special care should be taken to recognize factors that can be confounded with each other.

The detection data used in preparing this report can be made available in electronic form. Requests for the data should be directed to the Aging Aircraft Program, ACD-220, Federal Aviation Administration Technical Center, Atlantic City Int'l Airport, NJ 08405.

2. EXPERIMENTAL PLAN AND DESIGN.

The experimental hardware was designed and fabricated to simulate the fuselage of typical narrow body transport aircraft. Known flaws were engineered into two types of test specimens. Two monitors traveled with the experiment to set up the experimental hardware and to record inspection results. The choice of the lap splice and the curvature of radius that simulates the narrow body transports is one of convenience. The essential inspection characteristic being studied is that of detecting a crack originating within fastener holes in thin aluminum structure.

The experiment was taken to nine different facilities. The facilities were chosen to obtain a cross-section of those where inspections of transport aircraft are performed. Major attributes considered in the cross section included size of inspection force and in-house, versus third party inspection. At each facility, four inspectors (or inspection teams) completed the inspection task. At each facility, one of the four inspectors (or inspection teams) performed the inspection a second time. The net result was 45 inspections.

The intent of the experiment was to allow inspections to occur as they normally would in a non-test situation. Therefore, the equipment and the procedures followed were of the inspector's own choosing. The only potential departure from a routine inspection was in the way we asked the inspectors to record the results of their inspection. Crack detections were marked directly on a piece of protective tape that the monitors put into place before each inspection. The inspectors were also asked to give a subjective ranking (1, 2, or 3) reflecting their confidence that a flaw signal was present.

The need for the protective tape was verified in an experiment conducted during the planning phase. A typical pencil probe was rotated a number of times on bare aluminum as well as on various types of protective tape. From this experiment it was verified that without some form of protective surface a visual clue would be left at those sites where repeated inspections occurred. To keep the experimental conditions constant throughout the facility visits, it was deemed necessary to use protective tape. The tape selected was Scotch™ brand number 336. It is a transparent polyester tape with a very low tack rubber adhesive. Its total thickness is 0.0014 inch (0.04 mm).

It was determined through modeling, as well as empirically, that the presence of the tape should have minimal effect on the outcome of the inspections. The placement of protective tape on probe heads is not uncommon among inspectors. Through appropriate inspection setup procedures the presence of the protective tape is accounted for.

In section 2.1 the protocols used in taking the experiment to the various facilities are presented. The factors that were incorporated into the experiment are discussed in section 2.2. The information gathered while at each facility is discussed in section 2.3. The hardware and the flaw characteristics of the test specimens are discussed in detail in section 2.4.

2.1 EXPERIMENT PROTOCOL.

An extensive protocol was developed before visiting the various facilities that participated in the experiment. The protocol was developed to assure a consistency of operation from the first facility visited to the last. The protocol consists of a set of thirteen procedures. Here a brief description of each of the procedures contained in the full protocol is given. The full protocol is reproduced in appendix A.

RAE 1 *Instructions on Protocol Use.* This procedure describes how the Protocol was intended to be used.

RAE 2 *Inspector Supervision Procedure.* One purpose of this procedure is to describe the role and activities of the monitors in the field. Guidance is given on the type of comments required from the monitors on the inspector's performance, physical and mental skills, and personal characteristics relevant to inspection. It indicates how these subjective assessments can be quantified. The procedure also contains a summary of the actions required of the monitor in data recording and document control.

RAE 3 *Start Out.* This procedure controls the packing and movement of the equipment crates.

RAE 4 *Management Briefing.* This procedure is divided into two parts. To brief management on the aims and objectives of the PoD exercise, the monitor reads part A to them upon arrival at an inspection facility. The briefing session also allowed management to clarify any questions they may have had concerning the work. Part B contains a list of the points that the monitor must resolve with management.

RAE 5 *Start Up.* This procedure gives guidance on the operations required upon arrival at a new facility, covers discussions with management and setting up the equipment.

RAE 6 *Reference Standards Experiment.* This procedure details the steps involved in the Reference Standards Experiment. This experiment was designed to compare a facility's lap-joint calibration block(s) with the master reference block.

RAE 7 *Trial Checklist.* This procedure was used by the monitors to prepare for and record information specific to each inspection. It starts with the preparations for each inspection and takes the monitor through to the end of an inspection session.

RAE 8 *Panel Layout.* This procedure covered the setting up of the panel assemblies for each inspection.

RAE 9 *Inspector Briefing.* This procedure guided the monitor in informing the incoming inspector of the aims and objectives of the PoD exercise, and gave a general introduction on what the inspection involved and how it should be performed. The briefing was immediately before an inspection session.

RAE 10 *Pre-Trial Questionnaire.* Background information on the inspector(s) was gathered before the inspection of the test hardware.

RAE 11 *End of Trial Debriefing.* At the end of each eddy current inspection a record was made of each inspector's perception of how well his/her work went. This was achieved

by the monitor leading the inspector through the structured interview contained in this procedure.

RAE 12 *Data Recording and Transfer.* This procedure covers the recording and transfer of the inspection results from the marks on the test specimens made at the time of inspection.

RAE 13 *Close Down.* This procedure covers the operations involved in preparing to leave a facility at the end of all inspections.

2.2 CONTROLLED FACTORS.

In this section, the factors that were accommodated in the overall experimental design are presented. The rationale behind choosing these factors is also given. The factors are test specimen type, crack length distribution, off-angle cracks, inspection surface, accessibility, inspection time, shift work, crack density, and within-inspector repeatability. The manner in which each of these factors was addressed is described below.

2.2.1 Test Specimen Type.

Two types of specimens were fielded in this experiment. These are discussed in full in section 2.4.1. The 20-inch-square skin panels provided freedom to alter the presentation to each inspector. This minimized the transference of crack pattern knowledge from one inspector to the next. This was important because of the extended time (approximately one week) that the experiment was located at each participating facility.

There could be detection differences between artificial cracks assembled to final structure and real cracks produced by fatiguing final structure. Such differences were reported in Norriss [3]. Thus, the second type of specimen (full-size, large aircraft panels) was introduced into the experiment.

2.2.2 Crack Length Distribution.

The crack length distribution fielded in the experiment is presented in section 2.4.2. In evaluating the reliability of lap splice eddy current inspection procedures, Boeing Quality Control Research and Development Group concluded that "all procedures detected flaws 1/8 inch or longer with a 90 percent probability of detection and a 95 percent confidence" [4]. The experimental procedure used by Boeing employed in-house inspectors and did not reflect the myriad of field conditions that might impact the reliability numbers. It was believed that extending the flaw distribution to include flaws in the 0.2 inch range would be adequate to cover degradation that might occur in the field. This was also consistent with test results reported by Norriss [3] concerning eddy current inspection of outer skins.

2.2.3 Off-Angle Cracks.

Two levels (11° and 22°) of off-angle cracks were included in the lap skin specimens. The top level was chosen to reflect observed characteristics of field detected cracks [5]. Of the 122

fielded rivet sites containing cracks, 75 were horizontal, 21 were at 11°, and 26 were at 22°. A more complete description of crack characteristics is given in section 2.4.2.

2.2.4 Inspection Surface.

The intent in the design of the experiment was for inspections to be done as they would normally be performed. In a pre-visit questionnaire, the facilities indicated normal surface conditions, bare or painted, for their inspections. Some indicated that an inspector could be asked to perform a lap splice inspection on either type of surface. Based on the questionnaire responses, the surfaces of the 20-inch-square skin panels were bare at four of the sites. They were subsequently painted for the final five facilities. To enable a comparison of the effect of the painted condition that would not depend upon the facility choice, one of the full-size aircraft panels (the large panels) was painted and one was not. This condition was maintained at all the facilities.

2.2.5 Accessibility.

Intuitively, inspection reliability could deteriorate if the inspectors were forced into uncomfortable postures. To address this issue, one half of the skin specimens were presented at approximately knee height (24 inches), while the other half of the specimens were approximately five feet high. See figure 2.2. This presentation was done for every one of the inspections. The placement of skin panels was altered between each inspection at a given facility so that each crack was presented in the upper lap splice row twice and in the lower, knee high row twice.

2.2.6 Inspection Time.

The total amount of inspection included in this experiment was based on estimates of the amount of inspection necessary to take approximately four hours. This amount of time was thought to be enough to allow for the task to become routine, thereby allowing the inspector to "settle into" the inspection.

2.2.7 Shift Work.

Shift work was recognized as a potential variable that could influence inspection results. The original experiment design addressed this by requiring the inspectors at each facility to be chosen from each shift. The intent was to obtain approximately the same number of inspectors from each shift, over all the facilities. However, logistics in implementing the experiment required alterations in this plan.

Many of the facilities did not employ eddy current inspectors during a graveyard shift. Fewer inspectors worked the evening and graveyard shifts, thereby making it harder for the facility to free up inspectors in these shifts. Working with these types of constraints, seventeen of the forty-five inspections occurred on evening and graveyard shifts.

2.2.8 Crack Density.

Does the performance of an inspector change when many cracks are being detected versus when few are being detected? With this question in mind, each of the rows of lap splice was broken

into regions of various crack concentrations. Doing this also helped alter the perception of the inspectors to keep them from drawing conclusions concerning the amount of cracks present at any given time in the inspection.

In the skin specimens (see section 2.4.1 for specimen descriptions), each inspection row started with two unflawed panels. The purpose of having no flaws in the initial inspections was to counter the inspector's expectations for finding flaws. The remaining sixteen skin specimens were split evenly into an area containing approximately 10 percent flawed rivet sites and an area containing approximately 40 percent flawed sites. The order of these two areas within the inspection was altered so that half of the inspectors went from the high density to the low density during the inspection and the other half went from low to high. NOTE: The density of flaws includes smaller flaws (< 0.04 inch) that for the most part went undetected. Thus, the density of flawed sites observed by the inspector was less than that stated above.

2.2.9 Within Inspector Repeatability.

A given inspector using the same equipment and the same procedures may not achieve the same level of performance every inspection. Different results could result due to natural variations. To estimate how much variation could be expected in individual inspections, one inspector (or team) at each facility was asked to repeat the total inspection task. This was done by asking the inspector doing the first inspection to return after all four of the initial inspections had been completed. The order of the skin test specimens was changed between the two inspections so that the inspector could not rely on his recollection of flaw locations. The inspections also occurred at least three days apart.

2.3 SUPPORTING DATA.

Supporting data are those data recorded at each site for each inspection that were not a part of the controlled factors of the experimental plan. These supporting data were factors that could not be controlled in this experiment, such as light and temperature levels, or were not feasible to control, such as age of inspectors, previous training, and experience of inspectors, etc. Even though these factors could not be controlled in the experiment, they were recorded for the possibility of adding explanatory power to the results, as well as for suggesting possible influencing factors.

The supporting factors were not expected to have significant statistical power because of the lack of control over the balance and range of observed characteristics. For example, consider that one or two of the inspectors have a certain characteristic (e.g., female) and their inspection results are among the better results. Because of the low number of inspectors with the given characteristic, it is not unlikely that the results occurred by chance. That is, there would be no statistical significance to the observed inspection results as related to the given characteristic.

However, some of the observed supporting factors are analyzed (see section 4.5) for insight into reasons for the observed inspection results variation. The gathering of background data also allowed for a general assessment of the variations existing at inspection facilities. These observations are discussed in section 3.

We designated supporting factors in three major categories for recording: Environment, Equipment, and Personnel Characteristics (listed below). Some of the environmental factors could be objectively measured (e.g., lighting), while others required subjective judgments (e.g., housekeeping). The same was true of personnel factors, except that we relied upon the inspectors to provide such measurable factors as education, training, and age.

The monitor judgments and measures were recorded on Form RAE/7/Ins, while the self-reports were recorded on Forms RAE 10 and RAE 11. (These forms are part of the protocols and are included in appendix A.) Along with making the judgments required in these forms, the monitors reported general, overall impressions in separate trip reports that were abstracted from the monitors' personal logbooks. (The two monitors consisted of a Ph.D. Experimental Psychologist who is also a Certified Practicing Ergonomist and an ASNT Level II NDT Engineer in eddy current and ultrasonic testing.)

The following environmental factors were selected for measurement or recording on two bases. The first basis was their effects on inspection performance shown in other research. The second basis was the feasibility of measuring or consistently judging them. The selected factors were:

- Housekeeping
- Humidity
- Noise levels
- Changes in environmental conditions
- Temperature
- Lighting
- Tool condition and availability

The equipment factors that were recorded were the type, model numbers, and settings of the inspection equipment. The probe types and model numbers were also recorded. The inspection procedures followed and the scanning techniques employed were also recorded, as well as instrument calibration characteristics.

It was seen as easier to deal with the large number of personnel factors by breaking them into two categories: Physical and Psychological. Physical factors are more easily measured, judged, or reported (e.g., age or gender). Psychological factors are less easily measured, judged, or reported (e.g., attitude).

The method for selecting factors to be recorded in this experiment was similar to that for environmental factors, i.e., feasibility in consistently measuring or judging them. However, an additional criterion was necessary, the criterion of intrusiveness. The reason for inclusion of this criterion requires a brief digression.

It should be understood that the inspectors were regular employees of aircraft maintenance facilities. The parent organizations volunteered the time of these inspectors without any recompense. In fact, due to confidentiality requirements of the research, these facilities were not told how well their people did. That so many organizations volunteered the time and space at their own cost is a tribute to the dedication of these organizations to safety and the improvement of the entire industry. The inspectors were nominally volunteers -- although we know that most of them were simply assigned to the experiment in much the same way that they would be assigned to any other job.

The monitors carefully instructed the inspectors during the confidentiality briefing (RAE 9A in appendix A) that the inspectors could walk off the task at any time and the monitors would support that decision. In this way, we ensured that the inspectors were closer to true volunteers.

This system and approach meant that we needed the willing cooperation and hard work of the inspectors. That, in turn, meant that we could not ask questions that were likely to be perceived as offensive or intrusive. Some factors, like smoking or drinking habits, could be important indicators of physical condition and task performance. These types of questions were considered to be too intrusive and might destroy cooperation, so were not included. There were over sixty inspectors in the experiment -- none walked off, and only one called in sick on the day he was to repeat the experimental task.

Physical factors that were selected were:

- Observed general physical condition
- Age
- Amount of sleep during last off-period
- Prior activities
- Reported fatigue level (beginning)
- Postures used during test
- Gender
- Time on duty
- Reported physical condition
- Reported fatigue level (ending)

There were more psychological factors proposed overall, since the research literature is quite full of factors that have some effect upon performance (although some of those effects are quite subtle). However, it was not feasible to give these volunteers a full battery of psychological tests, so the factors were limited to those that could be obtained through self-report or that could be easily observed. In the design, however, some of these factors were obtained from both self-report and observation and, in some cases, several measures that get at the same factor were built into the system.

Psychological factors selected were:

- Education level
- General experience level
- Instrument-specific training
- Type of training
- Lap-joint experience
- Perceived management attitude
- Perceived effect of observers
- Reported attitude toward experiment
- Work patterns
- Attitude toward job
- Recency of instrument-specific training
- Instrument-specific experience
- Lap-joint experience recency
- Perceived realism of test
- Reported attitude during test
- Reported mental condition (e.g., irritability, efficiency, depression)

2.4 EXPERIMENT HARDWARE & FLAW STATISTICS.

2.4.1 Test Specimens.

Two types of test samples were used in the experiment. One type was 20- by 20-inch panels that could be moved and presented differently to each inspector to reflect factors of the experiment, discussed more fully in section 2.2. The second type of specimen was large panels that were produced with all the structural components found on an aircraft fuselage. Each is explained more fully below. The nature of the flaws in each type of specimen is discussed in section 2.4.2.

2.4.1.1 Lap Splice Joint (Skin) Test Specimens.

The 20- by 20-inch panels used in this experiment simulated fuselage lap splice joints found on the Boeing narrow-body aircraft. The specimens consist of two plates fastened together using three rows of rivets (figure 2.1). These specimens simulated the skins in the lap splice joint without any substructure. They were assembled on a frame and butted against each other to represent a longitudinal lap splice joint. Figure 2.2 shows two rows of skin specimens mounted on the frame.

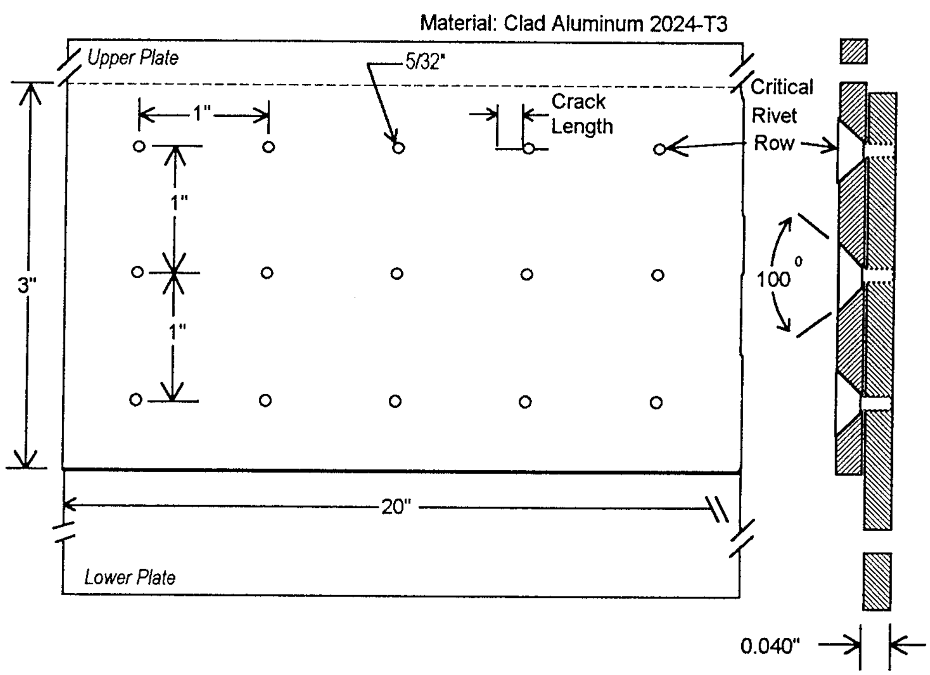

FIGURE 2.1 LAP SPLICE JOINT SKIN TEST SPECIMEN CONFIGURATION

2.4.1.2 Full-size Aircraft Panels.

Along with the lap splice skin specimens, larger test structures that simulated complete aircraft structure were used. These large panels (8 by 4 feet) were produced with all the structural components found on an aircraft fuselage. They contained one longitudinal lap splice joint and

were curved to match the nominal radius (75 inches) of the Boeing narrow-body aircraft. The panels were fabricated by Foster-Miller, Inc. and are shown mounted on the presentation frames in figure 2.3. Cracks were generated in the panels using a custom designed load machine that Foster-Miller developed for this purpose. The structural test frame provided a bi-axial load (hoop and axial load) that simulated the fuselage loads incurred during aircraft pressurization. The loads were applied in a cyclic manner and the cracks were allowed to initiate as they would in an aircraft [6].

The aircraft panels were designed to be more realistic than the skin specimens both in size and features. They were included in the experiment to assess any effects in the inspection task that could be attributed to the lack of total realism in the construction and crack placements in the skin specimens.

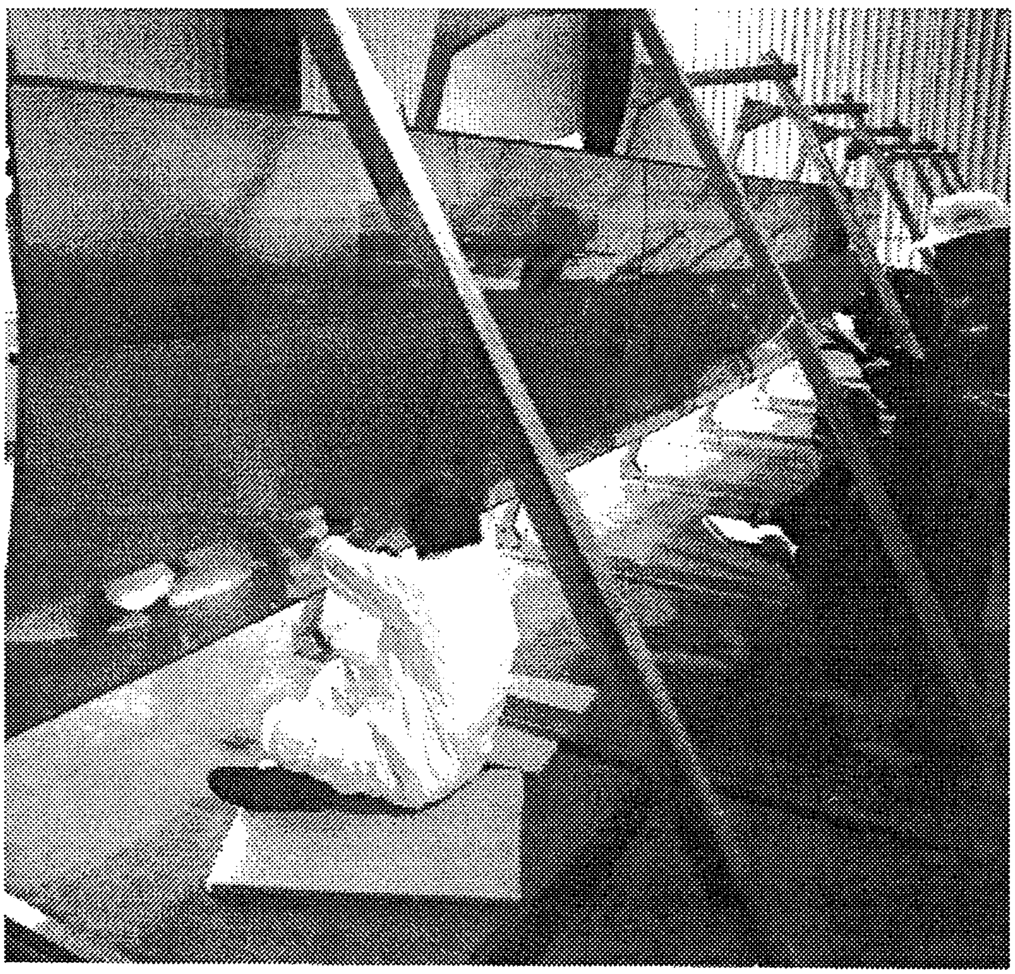

FIGURE 2.2 PRESENTATION OF LAP SPLICE SKIN SPECIMENS.
Inspector is shown inspecting bottom row.

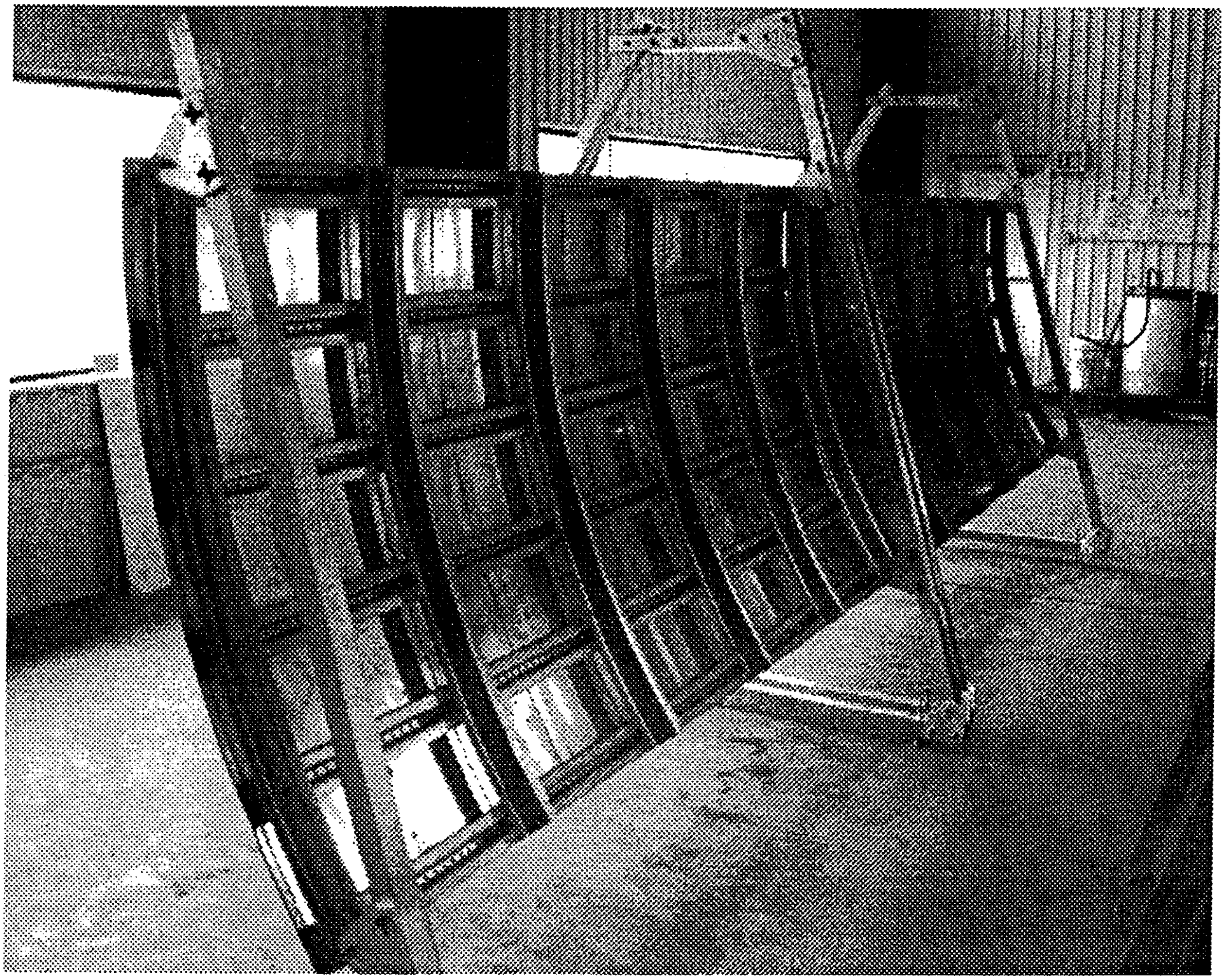

FIGURE 2.3 BACKSIDE OF AIRCRAFT PANELS SHOWING STRUCTURE.

2.4.2 Crack Characterization and Distributions.

2.4.2.1 Lap Splice Joint (Skin) Test Specimens.

For each of the skin specimen panels, flaw locations were specified and target sizes (lengths) were established. The locations that were to contain flaws were drilled out and notches were made to initiate the cracks. This was done in the upper plate only. The upper plate was placed in a mechanical test machine and a cyclic tensile load was applied to grow the cracks. Since fatigue cracks normally grow in the direction perpendicular to the load, off-angle cracks were produced by laying out the total upper sheet at an angle within a larger aluminum sheet placed on the test machine. Details of the generation of flaws are given in appendix A of reference 2.

Forty-three specimens were fabricated. Thirty-six of them were used in the experiment. The remaining 7 were built as backups in case of field damage. However, it was not necessary to employ them during the field experiment. At the time of the laboratory inspections it was not

known whether the backups would be used in the field. All test specimens were therefore included in the laboratory inspection data discussed in section 4.2.1.

Before the upper plate was assembled with the lower plate to make the completed lap splice, all rivet holes and associated countersinks were drilled. This drilling also eliminated the starter notches. The cracks in the upper plate were then measured on both sides by SAIC/Ultra Image, New London, CT. An optical comparator with a video-based microscopic viewing system and on-screen measurement tools was used (figure 2.4). The video micrometer (Oracle Model 6000) was capable of storing up to four calibration levels. Before the characterization, the operator calibrated the system at three different magnifications: 30X, 45X, and 60X. With the crack centered in the field of view, the calibration magnification levels were selected, and the level yielding the highest resolution (60X) was used for the crack measurement. A typical video image is shown in figure 2.5.

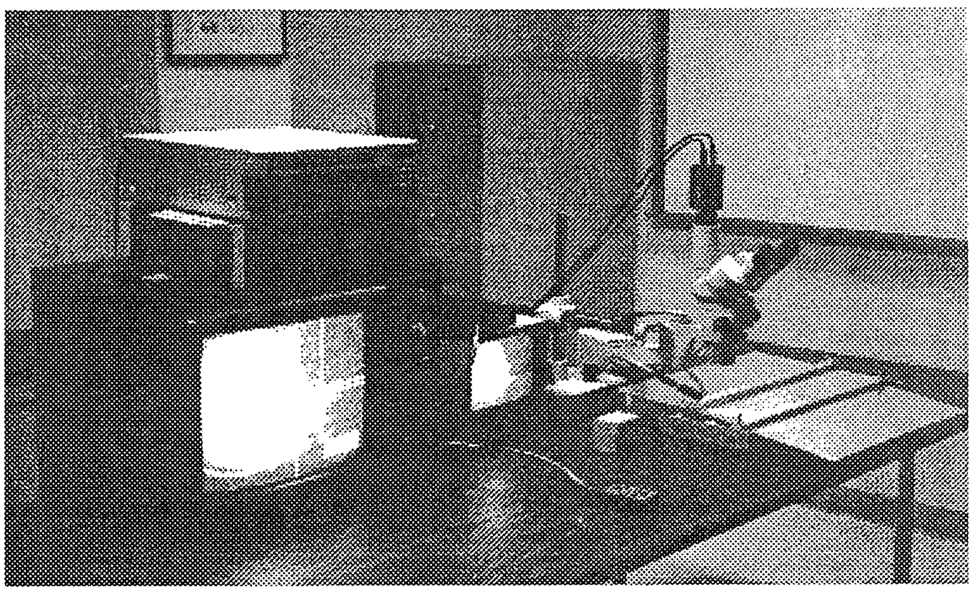

FIGURE 2.4 OPTICAL COMPARATOR SETUP USED TO MEASURE CRACK LENGTHS

Where measurements were available from both the front and the back sides of the upper plate, a given crack was characterized by the average of the measurements taken from each side. The average for all the cracks did not differ from a single side measurement by more than 0.008 inch. In 79 percent of the cracks with measurements on both sides, the average was no more than 0.003 inch different from the single-side measurements. Thus, we are assured that the cracks go through the thickness of the skin. Of the 215 cracks that were measured, in forty-two of them a measurement was obtained only on a single side. All of these were cracks within the countersink and all but two were measurements taken on the back side. In these cases, the single measurement was used to specify the crack length.

Upon completion of the characterization using the video measurement system, the skin panels were imaged on an eddy current imaging system consisting of an Ultra Image IV interfaced to a

Rohmann Elotest B2 eddy current system (figure 2.6). Images from cracks characterized from a single backside measurement were compared to images of similar size cracks where measurements were available from both sides. This was done to assure that the single back-side measurement was accurate when a clear measurement from the countersink side was not attainable.

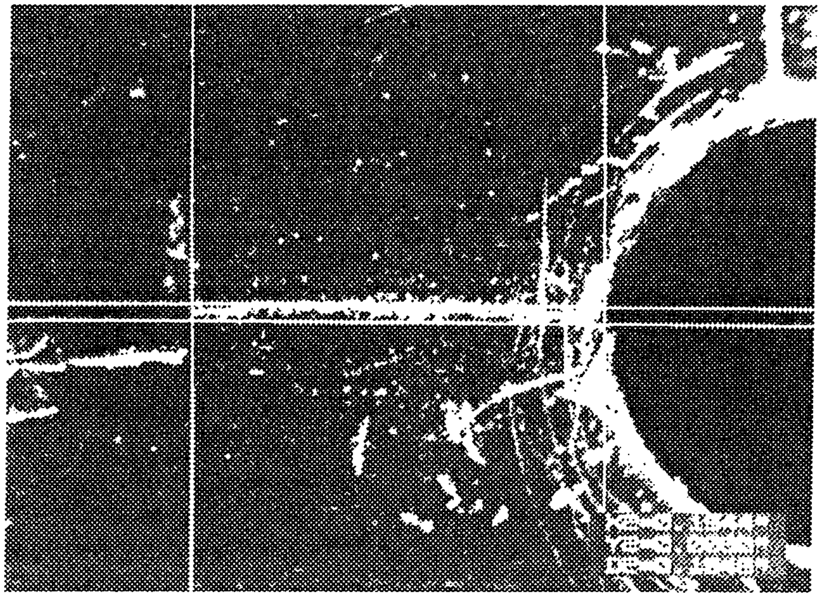

FIGURE 2.5 VIDEO IMAGE OF TYPICAL CRACK AT 60X.
Crosshairs show on-screen measurement capability.

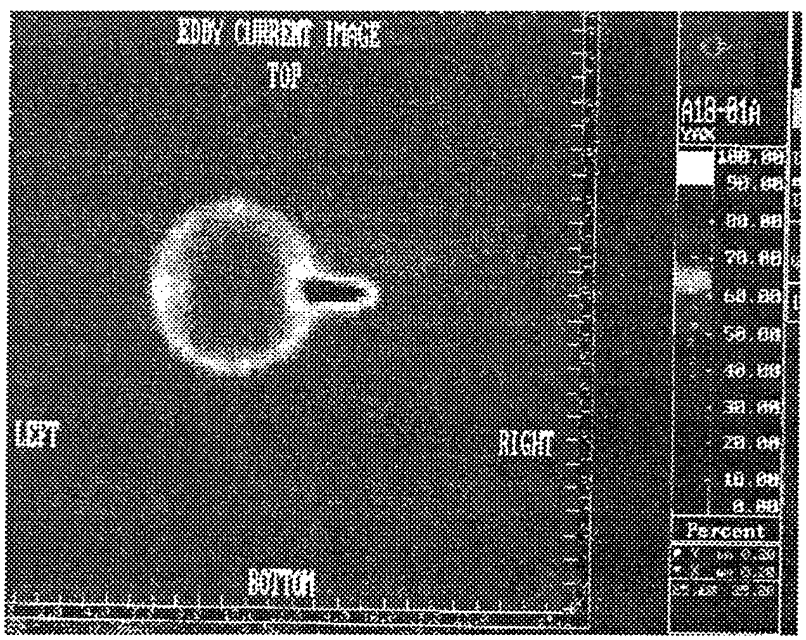

FIGURE 2.6 EDDY CURRENT IMAGE OF CRACK.

Some of the cracked rivet sites had cracks from the left side, some from the right side, and some from both sides. The cracks were horizontal, 11 degrees off horizontal, or 22 degrees off horizontal. The cracks that were off horizontal could be either above or below horizontal.

Because of the variability inherent in the crack growth process, it was not possible to obtain the exact crack length distributions as planned in reference 2. Table 2.1 summarizes the attained crack characteristics for the 20-inch lap splice skin specimens. For the off-horizontal cracks, the left (L) -- right (R) pairing in the table reflects the pairing that would occur within any given specimen. That is, if the left cracks were down from horizontal, then the right cracks would be up and vice versa.

Of the 172 cracks summarized in table 2.1, 100 of them occurred at rivets in which a second flaw was present. This information is summarized in the last row of table 2.1. The lengths of each of the cracks occurring at the same rivet site are shown in the graph of figure 2.7. There are 122 rivet sites in the skin specimens that contain cracks.

2.4.2.2 Full Size Aircraft Panels.

No target distributions were specified for the full size aircraft panels. The cracks were monitored and the fatigue cycling was halted when the largest crack reached about 0.20 inch. The video micrometer system used to characterize the skin panels was also used to measure and obtain hard copy images of the cracks on the full-size aircraft panels. However, because of the construction, with all the structure in place, only the outer surface could be imaged. The distribution of observed cracks is given in table 2.2.

Table 2.1 *Distribution of cracks by length and direction in skin specimens.*
( ) *denote distribution on panels used for spares.*

	Crack direction										
	horizontal		11 degrees				22 degrees				
			up	down	down	up	up	down	down	up	
Crack Length (inch)	L	R	L	R	L	R	L	R	L	R	Totals
0 to 0.020	2	1				1	(1)		(1)	1 (1)	5 (3)
0.020+ to 0.040	4 (2)	5 (1)	1	1 (2)	1 (2)	(1)		1 (1)	3	2 (1)	18 (10)
0.040+ to 0.060	7 (2)	7 (3)	1	3	(2)	1 (1)	1 (1)	1 (2)	2	4	27 (11)
0.060+ to 0.080	5	4		1	(1)		(1)	2 (1)			12 (3)
0.080+ to 0.100	9	8 (1)	3		2 (1)		1	1		1	25 (2)
0.100+ to 0.120	8 (1)	6 (1)	1		1		1 (1)	1 (1)	1	1 (1)	20 (5)
0.120+ to 0.140	3	5	1				(1)		1		10 (1)
0.140+ to 0.160	3	2					1	2		1	9
0.160+ to 0.180	3	3	(1)	1		1	1 (1)	(1)		1	10 (3)
0.180+ to 0.200	1 (1)	2	1	1					1		6 (1)
>0.200	11 (1)	9 (1)	1	1	2	1 (1)	2		2 (1)	1	30 (4)
Totals	56 (7)	52 (7)	9 (1)	8 (2)	6 (6)	4 (3)	7 (6)	8 (6)	10 (2)	12 (3)	172 (43)
second crack present	33 (4)	33 (4)	4 (1)	4 (1)	2 (2)	2 (2)	6 (4)	6 (4)	5 (2)	5 (2)	100 (26)

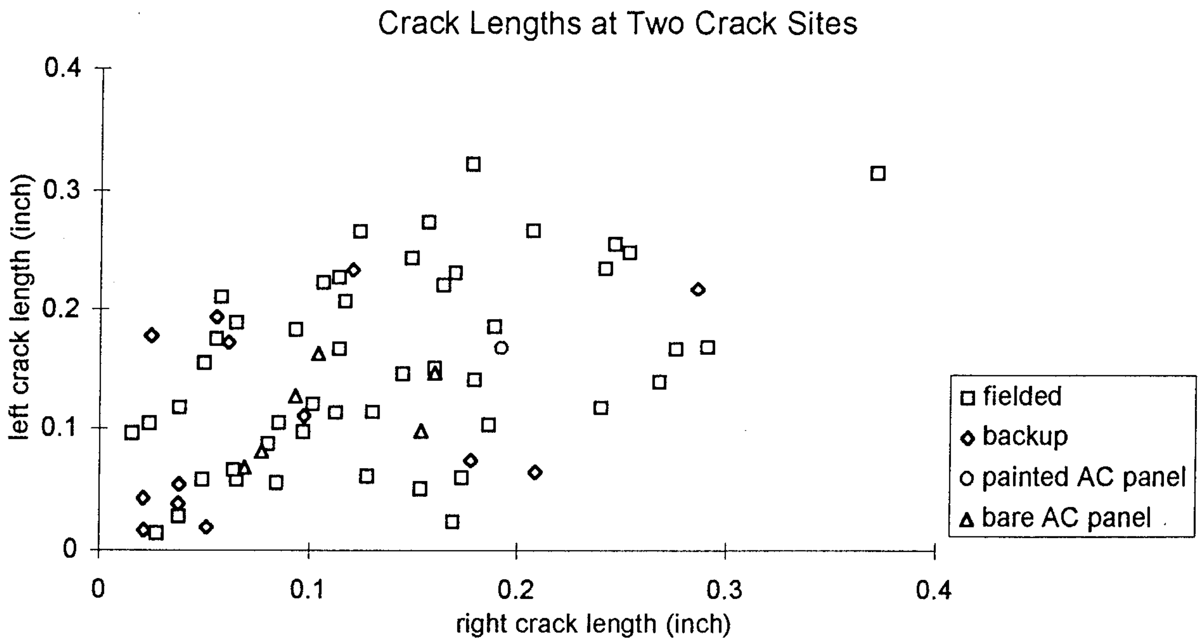

FIGURE 2.7 JOINT DISTRIBUTION OF CRACK LENGTHS AT SAME RIVET
Not shown are two sites at (0.812,0.506) and (0.378, 0.812).

Table 2.2 Distribution of known flaws in full-size aircraft panels.

	Painted Panel		Bare Panel		
Crack Length (inch)	L	R	L	R	Totals
0.060^+ to 0.080		3	2	2	7
0.080^+ to 0.100		11	5	1	17
0.100^+ to 0.120		11	2	2	15
0.120^+ to 0.140		8	1		9
0.140^+ to 0.160		7	2	3	12
0.160^+ to 0.180	1	4	2		7
0.180^+ to 0.200		1	1		2
Totals	1	45	15	8	69
Doubles	1	1	6	6	14

The two panels were constructed to the same specifications and in the same period. Note, however, that the flaw characteristics in each are different in both orientation and number. The painted panel received 47,435 load cycles and the unpainted panel received 66,533 load cycles.

For the aircraft panels, cracks beneath the rivet heads could not be seen by the optical system used to measure the cracks. As a result, all known cracks had lengths exceeding 0.060 inch. The inspection results, both in the laboratory and in the field, indicate that cracks are present that

were not verified optically. A complete characterization of the flaws could be obtained through sacrificing these specimens. However, it was decided that the specimens had more value intact (for future validation work) than would be gained from the characterization of the smaller cracks. This conclusion was reached primarily through the similarity between the results obtained for each type of specimen.

2.4.3 Post-Experiment Crack Characterization.

Following the completion of the experiment, a characterization of a sample of the skin panels was performed at the SAIC/Ultra Image New London, Connecticut facility. The post-characterization was performed to determine if further induced crack growth may have occurred due to shipping and handling experienced during the experiment. The characterization consisted of a repeat of the eddy current imaging that was performed in the initial characterization. The post-experiment samples were chosen to reflect the range of crack lengths.

The difference between the pre-experiment characterization and the post-experiment characterization is that the skin panels were fully assembled and painted during the post-experiment characterization. The effect of the rivet head on the eddy current signal is to exhibit a high amplitude signal around the edge. In the pre-experiment characterization, the countersink without the rivet installed exhibits a low amplitude signal. This difference presented a problem only with cracks that were within the countersink. In those cases, a trained and experienced inspector was relied on to determine the difference between the rivet head and a crack signal, or combination thereof, when watching the scope display.

Results indicate that no further crack growth had occurred on the sampled specimens during or after the experiment. Figure 2.8 shows the post-experiment eddy current image for the same rivet site as contained in figure 2.6. This is a typical example.

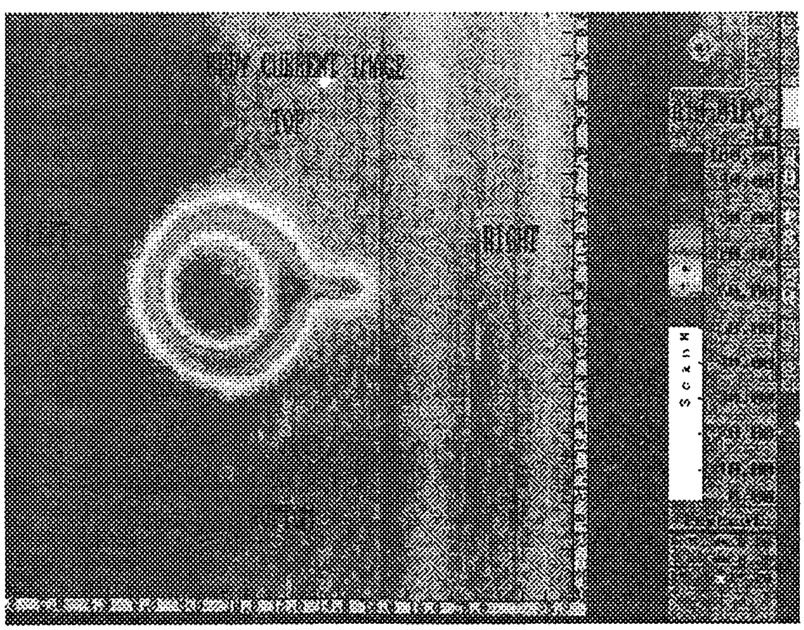

FIGURE 2.8 POST-CHARACTERIZATION EDDY CURRENT IMAGE OF CRACK.

3. CHARACTERISTICS OF FACILITIES AND INSPECTORS.

The eddy current reliability experiment was taken to nine facilities between April 1, 1993, and August 13, 1993. The facilities visited were American Airlines in Tulsa, Oklahoma; Dalfort Aviation in Dallas, Texas; Aloha Airlines in Honolulu, Hawaii; Tramco in Everett, Washington; Alaska Airlines in Seattle, Washington; United Airlines in San Francisco, California; Delta Airlines in Atlanta, Georgia; US Air in Winston-Salem, North Carolina; and Miami NDT in Opa-Locka, Florida.

The experiment was designed to answer questions related to eddy current inspections as they were being performed in the field. The cross section of facilities (large and small airlines, third party facilities, and so on) was chosen to represent the conditions under which typical inspections would be performed. Never were the individual facilities under test. To secure industry cooperation and to guarantee that individual inspection results could not be used for or against the facility or the inspectors within the facility, confidentiality of the results was promised and has been maintained. In keeping with the confidentiality, the following discussions do not specifically identify any one facility with the data.

In the following sections, observed facility and inspector characteristics are presented. These are presented to illustrate existing variations and to provide a contextual background for the data gathered.

3.1 FACILITY BACKGROUNDS.

The experimental program was taken to four major air carrier maintenance facilities, two small carrier facilities, and three independent contract maintenance facilities. Six of these facilities were union shops; however, the inspectors were not always union members -- depending upon whether the contract considered them management or labor. Four facilities were located at very busy airports. The rest were located at smaller, more regional airports. The general conditions found at each facility are briefly discussed.

The experiment was set up in hangars for eight of the nine facilities. For the one exception, the experiment was set up next to the sheet metal shop in an unused corner of the shop facilities. In all cases the inspectors were exposed to the same type of conditions and distractions that would be present during an actual aircraft inspection.

- Housekeeping. Housekeeping was noted because of the direct effect that housekeeping in the environment has been found to have on performance through attitude, morale, and motivation. Moreover, this factor has been found to be a useful and meaningful indicator of management attitudes toward employees and towards business in general. See Crosby [7] and Deming [8].

  Housekeeping in the facilities varied from extremely clean to quite dirty. Some of the facilities were nearly spotless, most showed that they were maintenance facilities, but there were efforts to keep loose trash swept up. A few facilities had trash on the floor (a thick

layer of dirt and odd bits of metal, screws, and nuts) for the entire several days we were there.

Facilities also varied from spacious and well organized to extremely cluttered and crowded. Cleanliness was correlated with crowding and clutter, since it is more difficult to keep cluttered areas swept and picked up.

- Temperature and humidity. Our humidity meter seemed to read about 20 percent low, compared to the reported humidities on the newscasts. This low reading seemed quite consistent. One facility had an enclosed and temperature controlled environment. The other eight were open to the ambient environment, although portable fans and heaters were sometimes available. One facility did not have closable hangar doors, although it was in a very even environment.

- Lighting. Lighting varied from a very usable 100 foot-candles (ft-c) to less than 10 ft-c, which was the lowest our instruments could measure.

- Noise levels. Noise levels varied from a very quiet, office-like 55 dB to extremely noisy and potentially ear-damaging 110 dB. Since the facilities were all located near runways, even the quiet facilities were subjected to sudden noise peaks of 85-120 dB (depending on the facility).

- Tool condition and availability. On the whole, tools were available and adequate in all facilities. Smaller facilities tended to have older instruments.

- Changes in environmental conditions. Only one facility did not have wide changes in environmental conditions during a shift. This facility provided plentiful lighting and good climate control. The hangars at this facility were large enough to dampen noise levels a great degree. The rest of the facilities that were visited had environmental conditions that changed with time of day, with changes in ongoing activities, and with location on the aircraft where inspection occurred (especially lighting and noise conditions).

  Day shift had more uniform temperature and better lighting (both factors tended to change more if hangar doors were opened). Day shift also had more noise, both from other work and from nearby aircraft. Evening shift usually started with good lighting and high noise, but both lighting and noise tended to decrease after 7 or 8 p.m. Night shifts were quiet with few interruptions, but lighting was often problematic. During the experiment, there were occasions when coats were needed, but gloves were never needed. In the deep winter, that situation might change slightly, but for the most part, heating was reasonably good; air conditioning was poor at all but one facility.

  The experimental inspections were always inside, but the monitors reviewed conditions outside to note the differences. The major difference was lighting and glare. Both had higher values outside, including at night, since large auxiliary lighting was usually available. Environmental differences would be more pronounced outside, with wider temperature swings and more wind and noise to contend with.

- Equipment. Equipment ran the gamut from older Magnaflux ED-520s and Forester 2.8s to new Zetec MIZ-22s and Nortec 19s (and one Rohmann B1 with a rotating surface probe). All the inspectors in seven of the nine facilities used the same type of equipment. Two and three different types of equipment were used at the remaining two facilities. The equipment is tabled in section 4.2.2.

- Procedures. The Boeing procedure [9] describes three different eddy current inspection methods designed to detect cracks that extend beyond the edge of the countersink. The three methods are the sliding probe, oversize template, and rotating surface probe. The three methods are optional to each other, but both the sliding probe and the rotating probe call for verification with the oversize template method.

  All facilities had Boeing 737/727 inspection procedures available; however, some facilities had adapted the Boeing procedures for their particular organization. There were variations in the ways that those procedures were used and interpreted. Several inspectors brought the procedures with them but did not look at them. Several inspectors did not bring the procedures with them; although many mentioned they had reviewed them when they heard that they would be in the experiment.

- Instrument calibrations. There was wide variation in the frequencies and settings that were used in setting up the instruments. There also were several different kinds of standards blocks used; although the Boeing #290 block was most common. In two facilities nonstandard calibration blocks (i.e., not in the Boeing procedures) were used. In these facilities the set up of the equipment was accomplished on electro-discharge machined (EDM) notches on "universal" eddy current standards. When asked, the inspectors stated that the setup on the nonstandard blocks would produce the same results as setups performed on the Boeing #290 or #369 blocks. In one facility, the inspectors did not use a calibration block at all; instead calibration was done on a supposed crack-free rivet site on the test specimens.

  There was variation in the calibration procedures beyond the choice of calibration blocks. For example, some inspectors calibrated impedance-plane instruments to have a very tall, narrow loop; while others calibrated for a very short, flat loop. Similarly, some inspectors calibrated needle-deflection instruments to small deflections for lift-off and large deflections for cracks (typically full-scale); while others calibrated to deflection in one direction for lift-off and in the opposite direction for cracks.

- Scanning techniques. Ten of the 22 inspections using the sliding probe did not use a straight edge at all or discontinued the use of a straight edge relatively quickly into the task. When questioned, inspectors responded that most rivet lines were too uneven on most aircraft for a straight edge to be useful. Inspectors tended to use sliding probes from left to right, then reversing the direction for a second check. Inspectors stated that the sliding probe was more sensitive in the direction of travel, and that the "backscan" was a double check.

A very few of the inspectors who used a pencil probe did their scans "freehand," without a template -- although freehand methods were frequently used when there was a question of a small crack near the edge of the rivet head. Only one inspector used double-sided tape on the hole template. All other inspectors held the template with one hand, while scanning with the other hand.

One team was found to use the sliding probe without following up with a pencil probe and template. They had found that they could determine the orientation of a crack by noting whether the flying dot made the loop in a clockwise or counterclockwise direction.

3.2 INSPECTOR BACKGROUNDS.

Data on the following physical factors connected with the inspectors were gathered from questionnaires or were observed directly by the monitors.

- Observed general physical condition. Most of the inspectors were in average physical condition for American shop-floor workers. There was one obese, stiff inspector and four or five athletic inspectors. A few (four or five) inspectors reported feeling unwell (one was succumbing to flu, others had headaches or allergic reactions).

- Postures used during test. All but one inspector stood while inspecting the top row of small panels. The exception was an inspector so short that he/she had to stand on a Coke crate, or kneel on a small shipping crate to reach the top row. Positions taken for the lower row of small panels varied much more. Some inspectors lay on foam pads or an automotive-style creeper (where available). Some sat on foam pads or the creeper. One location had a special low chair on rollers that the inspectors used. Two inspectors took martial-arts wide squat stances through the full row.

- Age. The average age of inspectors was in the 31 - 50 range; the majority of inspectors were in this age range. Some inspectors were as young as their early 20's and the oldest in their mid-60's. At most locations, NDT inspectors are required to have experience in sheet metal or A & P mechanics before being hired or moved into the NDT department; therefore, the strong majority of them are over 30 years of age.

- Sex. Only two of the inspectors that participated were female; although there were two or three more female inspectors in the facilities that were visited.

- Previous amount of sleep. Most inspectors reported sleeping well or very well, but nine said that they had slept poorly to very badly prior to the inspection. With one exception (an individual who was having family problems), all inspectors who reported poor sleep were working the evening or the midnight shift. All of these also reported having been awake for some time and involved in family or personal business activities.

- Time on duty. This experiment took almost a full shift, including the time required to brief inspectors as well as to administer the pre- and post-questionnaires. Two inspectors were on the second of a double shift and had worked for the previous eight hours performing

aircraft inspections. The rest of the inspectors began the inspections within two hours of starting their shift.

- Prior activities. This question was designed to quantify the fatigue problems that we might find if the inspector was in the second half of the shift. However, the session took an entire shift and the question had relevance only for double shift inspectors as mentioned above.

- Reported physical condition. On the average, the inspectors reported feeling "OK", that is, about normal. The monitors observed that most of the inspectors were energetic and seemed healthy to very healthy with three exceptions, two double-shifting inspectors and one who was coming down with influenza (and called in sick for the next several days). There were some inspectors with mental distress or distractions; those are discussed in the discussion of "Psychological Factors." One female inspector was pregnant.

- Reported fatigue level (Beginning). Those inspectors reporting that they felt mentally or physically tired were mostly those on late shifts. Only one team and one person not on a team reported any tiredness at the beginning of the session. The solo inspector reporting tiredness was having family problems and disturbed sleep.

- Reported fatigue level (Ending). About half the inspectors reported being tired at the end of the experimental session. Inspectors' comments indicated that they spent more time actually inspecting and less time socializing, coping with paperwork interruptions, etc., than they would normally do during a shift. Several inspectors commented that the length of the lap splice used in the experiment (77 feet) was more than they would expect to do in a single shift.

Data on the following psychological factors were gathered from questionnaires or were observed directly by the monitors.

- Attentiveness. Both self-reports and monitor observations were used to collect data on attentiveness. Two inspectors said they were bored with the whole thing from the beginning. Two inspection teams expressed boredom when they had to repeat the session. Two more inspectors answered "yes and no" when asked if they had found the experimental inspection interesting. Only one of the inspectors who reported boredom, however, allowed his gaze to wander while working.

- Observed attitude. Monitors marked 5-point scales for the behaviors: interested, cooperative, hardworking, motivated, careful, and conscientious. Although subjective on the part of the monitors, the intent was to provide a relative ranking of the overall perceived attitudes of the inspectors. Inspectors in this experiment were above the midpoint on all scales with two exceptions. These two inspectors were from different sites and were working different shifts, so there was no common attitude or fatigue factors that would explain their attitudes.

- Work patterns. Work patterns varied among inspectors and facilities. In some facilities, inspectors were careful to take breaks and did not want to work over the end of the shift.

Some inspectors hurried the last 40 or 50 inspection sites to finish on time. In other shops, inspectors took their time and did not worry about exceeding the end of the shift by an hour or so. Some of the inspectors who took fewer and shorter breaks during the inspection were also called off the task for other duties more often.

There were no other work patterns such as conscientiousness, carefulness, precision, etc., that were characteristic of any one facility or shift.

- Education level. All inspectors had high-school diplomas (or equivalent). Most of them had attended college or technical school and most had A & P licenses. Many inspectors had NDT training from military service.

- Attitude toward job. Most of the inspectors participating in the study expressed a liking for inspection and wanted to do a good job at it. Layoffs were being discussed in one form or another in four or five of the facilities.

- General experience level. Most of the highly experienced inspectors were in the larger airlines. The contract maintenance facilities appear to have a high turnover, leading to many more inexperienced (but not necessarily bad) inspectors. One or two of the most experienced inspectors got the worst performance scores, however.

- Instrument-specific training and experience. Experience ranged from inspectors who had only looked at the manual and then started using the instrument to inspectors with intensive training on the instrument that they were using. All the inspectors in this study had some experience with the instruments they were using. Most had been using eddy current equipment daily or several times a week.

- Type of training. About 75 percent of the inspectors in this study had received some sort of classroom training on nondestructive testing (NDT), eddy-current instruments in general and/or on their specific instrument. Of these 75 percent, about one-third had only had local facility training. Other training locations included instrument manufacturers' training, Boeing training, and Hellier & Associates training.

- Lap-joint experience. Most of the inspectors had performed lap-joint inspections frequently when the Airworthiness Directives (ADs) regarding lap splices were first issued. However, most of the participating airlines had taken the required remedial steps or gotten rid of their 727s and 737s where the problem might be found. No lap splice joint inspections had been done for many months by inspectors at the airline facilities. Inspectors at contract maintenance facilities had performed lap splice joint inspections more recently, since they inspect foreign aircraft not having completed the remedial action.

- Inspectors' perception of management attitude. When asked what they thought their management's priority of nondestructive inspection (NDI) was compared to other things, twenty-five of forty-eight inspectors reported that they thought their management placed high or top priority on NDI work. However, in two of the facilities the inspectors were unanimous in reporting that they thought their management placed a low priority on the

importance of NDI. In a third facility, three of the four inspectors expressed an opinion that the NDI priority was high within the QA organization, but low in the maintenance organization overall.

- Perceived realism of test. All inspectors rated the test panels at the category "similar in some ways, not in others" or at the category "Very Similar." The major dissimilarities between the test setup of the study and actual aircraft, according to the inspectors, were:

    (a) They would never have to work on something in the position of the lower row of panels. They would always use scaffolding, ladders, cherry pickers, or something to improve access.

    (b) The test panels were smoother, more uniform, with rivets better aligned and more consistent than usually found on real aircraft.

- Perceived effect of observers. Only 5 inspectors (less than 10 percent) said that they were bothered "Some" by the watchers. All other inspectors said that the watchers did not bother them at all. Many said that they were very accustomed to watchers standing around when they worked. Others pointed out that they usually do inspections (other than radiography) with other people working all around them.

- Reported attitude during test. Over 90 percent of the inspectors were very cooperative and wanted to help get good data for this experiment. Even those who felt that they had other work that they needed to do, felt that the test was important and meaningful and wanted to do their best.

    The approximately 10 percent who were less cooperative included one or two who had personal problems that they admitted were occupying their minds. Two of the inspectors expressed a dislike to the overall experimental situation.

- Reported mental condition (e.g., irritability, efficiency, depression, mental condition). Most of the inspectors rated themselves "about normal" on these conditions, with the exceptions of one or two with problems in their personal lives that depressed or distracted them. One inspector who rated himself low on thinking ability and alertness had just finished one 8-hour shift and was starting his second one.

- Reported attitude toward experiment. Nearly all the inspectors felt positively toward the experiment. Many were unhappy that they could not find out how they scored, however. Almost all the inspectors and many of the NDT supervisors expressed hopes that the experiment would gain the attention of upper management regarding the importance of NDT inspection and NDT departments. There were also expressions of hope that the experiment would result in uniform standards of inspection and inspector qualification.

4. STATISTICAL ANALYSIS OF DETECTION DATA.

In this section we review the statistical methods used in analyzing the detection data and present the results of that analysis. The review of the analysis methods is information for the interested reader but is not meant to be a tutorial on statistical procedures. References are given for the reader who is interested in pursuing the methodologies in more detail.

An attempt has been made to state the method of analysis, present summary graphs and charts, and state the results, but relegate the statistical detail to appendices.

4.1 REVIEW OF ANALYSIS TECHNIQUES.

As indicated by this report's reference list, much has been written concerning the probabilistic models and the statistical procedures used in characterizing NDI equipment and procedures. In section 4.1.1 we review the common methods used for fitting probability of detection curves to binary (detect or no-detect) data. The estimated model parameters (or specific points on the PoD curves) for each inspection can be analyzed with respect to other potential explanatory factors. This approach is presented in section 4.1.2.

Probability of detection models are built upon data gathered concerning the detection or lack of detection of flaws. Data on the false calls are not used in estimating PoD curves. Relative Operating Characteristic (ROC) curves address the issue of total accuracy (in calls made on both flawed and unflawed sites) by incorporating false-call data into the analysis. An overview of this procedure is given in section 4.1.3.

Data analyses specific to the Eddy Current Inspection Reliability Experiment are presented in sections 4.2 to 4.5.

4.1.1 Regression Analyses for Probability of Detection Curve Fits.

In this section a brief review of binary regression models for hit/miss (detect/no-detect) is given. The binary or "hit/miss" refers to the use of NDI procedures to detect only the presence or absence of flaws. No additional information about flaw characteristics is given by the procedures.

The Boeing procedure [9] calls for an instrument calibration step in which a reference standard with a known flaw is used to ready the equipment. For the template and rotating probe methods the procedures state that a call should be made for any signal similar to the reference calibration signal. For the sliding probe method the criterion for a call is a one division shift in the signal when the setup was to have at least a two division shift between the non-defect signal and the defect signal. For all three methods, the inspector ultimately has to make a yes/no or binary decision.

Probability of detection curves have been used extensively to assess the accuracy or reliability of NDI systems and procedures. Background discussion of PoD curves can be found in Berens [10], Annis, et. al. [11], and Hovey and Berens [12].

The general form for PoD curves can be expressed as follows. Define the random variable Y as follows

$$Y = \begin{cases} 1 & \text{if site is judged flawed} \\ 0 & \text{if site is judged unflawed,} \end{cases}$$

and a is the crack length. The probability of detection can then be expressed, for example, as

$$\mathrm{PoD}(a) = \Pr(Y{=}1 \mid a) = F(\alpha + \beta \bullet log(a)), \qquad (1)$$

where α and β are parameters to be fitted to the data, F is a cumulative distribution function, and *log*( ) is the logarithm function. (In general, the argument of F is an increasing function of a -- possibly different from the log-linear form given here, which is most often used.)

We present three choices for the distribution function F. The choices presented include the two most common forms of PoD curves. They are the logistic (log odds or logit) and the normal (probit). We also briefly discuss a third, the Gompertz distribution because of its connection to the Weibull distribution, which has been used in modeling PoD curves.

$$\int_{-\infty}^{x} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \qquad \text{(normal)}$$

$$1/(1+\exp(-x)) \qquad \text{(logistic)}$$

$$1\text{-}\exp(-\exp(x)) \qquad \text{(Gompertz).}$$

Letting $\alpha = -c \bullet log(b)$ and $\beta = c$, then the Gompertz distribution evaluated at $\alpha + \beta \bullet log(a)$ becomes $1\text{-}\exp(-(a/b)^c)$, which is the two parameter Weibull distribution.

The general model given in equation (1) uses the logarithm of crack size as the explanatory factor. These models are sometimes developed with the crack size, a, rather than with the logarithm of a. To distinguish between the two, we will refer to lognormal, log logistic and Weibull forms of the PoD for the three distribution functions when $log(a)$ is used. The use of $log(a)$ guarantees that PoD(0) = 0, whereas the use of a often implies PoD(0) is positive. This may be of little practical significance as $F(\alpha)$, for the fitted α, may be arbitrarily small.

Berens and Hovey [13] evaluate seven different functional forms for PoD. They determined that the log logistic was the best among those tested for their NDI application. It has also been shown [10] that the lognormal and log logistic distributions produce very similar curves. This is also true for the data gathered in this experiment.

The basic PoD model can be extended to include explanatory factors other than $log(a)$ (or a). This is done by expanding $\alpha + \beta \bullet log(a)$ into more parameters that denote the state of other factors present at the time of inspection. For example, instead of modeling the inspection only with the two parameters, α and β, we might consider using $\alpha_i + \delta_j + \gamma_k + \beta_{ijk} \cdot ln(a_{ijkl})$, where i = 1 or 2, according to whether the inspection surface was painted or was bare, j = 1, 2, or 3

according to whether the inspection occurred on the day, evening, or graveyard shift, and k = 1 or 2, according to whether the inspection occurred on the upper or lower row of lap splice. In essence, by fitting this extended model we would be fitting different PoD curves to the different conditions under which inspections are done. The statistical procedures that provide for the estimates in such models also provide for an assessment whether the fitted parameters are significantly different from each other.

The second type of extension of the basic PoD model is to add a threshold parameter. The addition of a threshold is very common in medical applications. For example, consider quantifying the likelihood of contracting cancer as a function of a dose level of some cancer-causing agent. The use of a threshold allows for the naturally occurring background cancer rate to become part of the model. In the application to PoD curves the assumption is that there could be a background miss rate that is completely independent of crack size and due more to procedural errors. The need for such a model was suggested by observing inspectors and noting certain habits that would lead to missed cracks regardless of the length of those cracks.

Specifically, several of the inspectors were observed stopping their inspection to move equipment and then resuming their inspection at a different point, thereby missing several rivet sites. Inspectors also reacted to distractions (e.g., loud noise, conversations with other personnel) and it was not clear that they had maintained their attention through a completed inspection at a given rivet site. One inspector relying on an audible alarm did not immediately respond to hangar ambient noise rising to levels that could effectively mask the instrument alarm. Some of the inspectors experienced intermittent problems with their equipment. All of these conditions lead naturally to considering a PoD model that incorporates a background miss rate independent of crack size.

This extension of the PoD model is accomplished by replacing the model of equation (1) with

$$\mathrm{PoD}(a) = \Pr(Y=1 \mid a) = (1-C)\bullet F(\alpha + \beta\bullet log(a)), \tag{2}$$

where C is the background miss rate. The parameter C is estimated along with α and β from the data. Note, that the effect of adding this background miss rate is that of having an asymptote other than 1 for the probability of detection, given arbitrarily large cracks. This model is discussed in more detail in section 4.2 when individual inspector's results are presented.

4.1.2 Regressions on Individual PoD Parameters.

The basic motivation for adding possible explanatory factors discussed in section 4.1.1 was to be able to alter the coefficients, α and β, according to different conditions under which inspections occur. Achieving this sort of separation requires an adequate amount of data. In this experiment all the inspectors inspected the same set of test specimens, which contained enough flaws over a broad range of crack lengths to make it possible to fit tight PoD curves to each inspection.

Having obtained an α and a β for each inspection, we can consider them as responses and determine if their variation can be explained by other factors, such as the background experience

of the inspectors or the procedures used by the inspectors. The statistical model would be $\alpha = \overline{\alpha} + E_i + P_j + e_{ij}$, where $\overline{\alpha}$ is the average level of α, and i indexes levels of experience, j indexes the procedures used. Analysis of variance (ANOVA) is the statistical technique that can be used to ascertain whether certain factors are contributing to the observed variation in the responses. Examples of this technique and its extension to a joint (or simultaneous) analysis of two response values together (Multivariate analysis of variance or MANOVA) are given in appendix D of reference 11.

The ANOVA and MANOVA approaches are generally more meaningful when applied to functions of α and β rather than to α and β directly. To appreciate why, let the lognormal distribution be used to describe the PoD. Then the PoD curve is easily derived from the lognormal distribution with mean, $\mu = -\alpha/\beta$ and standard deviation, $\sigma = 1/\beta$. Let a_x denote the crack length for which $PoD(a_x) = x/100$. Then $log(a_x)$ corresponds to the x percentile of the lognormal distribution with parameters μ and σ. Thus, $log(a_{50}) = \mu$ and $log(a_{90}) = \mu + (1.2816)\bullet\sigma$. In general, $log(a_x) = \mu + (z_x)\bullet\sigma$, where z_x is the x/100 percentile from the standard normal distribution. Selected percentiles are generally of more interest and utility in characterizing inspection performance than are either of the two model parameters α and β.

Natural candidates to consider as responses for ANOVA analysis are $\hat{\mu}$, $\hat{\sigma}$, or $\hat{\mu} + (z_x)\bullet\hat{\sigma}$ for values of z_x needed to obtain selected percentiles. These approaches are taken in section 4.5 for the eddy current inspection data, where the responses are $\hat{\mu}$ and $\hat{\mu} + (1.2816)\bullet\hat{\sigma}$, which estimate $log(a_{50})$ and $log(a_{90})$.

4.1.3 Relative Operating Characteristic (ROC) Curves.

The PoD characterization of inspection reliability only considers one aspect of the inspection process: that is, the possibility of missing flaws that are present. Although this is extremely important from the aspect of safety, it ignores an aspect of reliability that is important from an operational point of view. The probability of calling an unflawed area as flawed is the other side of the coin that is not been considered when looking only at PoD curves.

For any inspection system the probability of a false call (PoFC) and the probability of detecting a true flaw are related. At one extreme -- that of calling a flaw for every inspection -- PoD would be a desirable 1, but PoFC also equals 1, which is undesirable. At the other extreme -- making no calls -- PoFC would be at the desirable 0, but PoD would also be 0.

The Relative Operating Characteristic (ROC) curve gives the relationship between the two probabilities. The ROC curve is a plot of the probability of detecting a true flaw (PoD) against the probability of a false call (PoFC) as the decision-making threshold is varied. ROC curves pass through the points (0,0) and (1,1) in keeping with the above discussion.

An ROC curve is fundamentally different from the PoD curves already discussed. There is no explicit use of crack length as an explanatory variable. The relationship between PoD and PoFC

reflected by the ROC curve is due to changing the criterion (or amount of evidence) that is required to judge that a flaw is present. A single ROC curve can reflect a distribution of flaws (or flaw sizes) or one can plot a different ROC curve for each specific set of flaw characteristics.

The use of ROC curves has been extensive in medicine and psychology for evaluating the effectiveness of diagnostic systems. Swets and Pickett [14] provide a summary and justification for the approach. Swets also discussed the use of ROCs and signal detection theory in NDI applications [15, 16]. Applications to NDI can be found in Davis [17] and Glasch [18], who used ROC curves to evaluate operator reliability.

Changes in the criterion (or decision threshold) for making a call have to be made to obtain different points on any given ROC curve. Different decision levels can be obtained by asking the inspector to use a subjective rating for each positive call (a flaw is present) that is made. One point on the ROC curve would be the proportion of detections versus the proportion of false calls where a positive call is considered to have been made for only the highest (most certain) rated calls. A second point would be obtained by considering as a positive call any calls made with the top two ratings. This process would continue, each time adding the next rating level. This is the approach used here.

4.2 POD INSPECTION RESULTS SUMMARY.

The Probability of Detection curve fits for the data gathered in this experiment were obtained with the commercially available software SAS®. Specifically, the SAS procedure PROBIT was used. The PROBIT procedure calculates maximum-likelihood estimates for regression parameters and for threshold response rates (if included in the model). Details are given in reference 19.

4.2.1 Laboratory Inspections.

Before taking the experiment to the various facilities, we had four inspectors inspect the specimens in the laboratory. There were no constraints placed on the inspections other than the use of currently available eddy current equipment. The purpose of these inspections was to establish a baseline that could be considered as reflective of the detection capabilities in the best of environments. The background of the inspectors and the procedures used are given in table 4.1.

***Table 4.1** Background on inspectors and procedures for laboratory inspections*
All used Rohmann Elotest B2 instrument.

Inspector	Qualifications	Procedure	False Call Rate (%)
EC1	NDT Level III, Supervisor and Examiner - >25 years aircraft experience	Sliding probe - 16 kHz	0.9
EC2	NDT Level III, ASNT Level III - >25 years in NDT training	Template - 20 kHz	0.3
EC3	ASNT Level III - >18 years in NDT training and development	Template - 30 kHz	0.3
EC4	ASNT Level II - 13 years experience in NDT inspection and development	Template - 30 kHz	0.9

The various PoD fits for each laboratory inspection are shown in figures 4.1 to 4.4. The similarity between the log logistic and the lognormal curves is apparent for all four inspections. The Weibull shows a little more departure from the other two, although in general exhibiting the same behavior. The biggest departure of the Weibull from the other distributions is in fitting the data of inspector EC4. Inspector EC4 missed two flaws with lengths of 0.105 inch and 0.108 inch. The effect of the misses at the larger cracks is to "flatten" the PoD curves. However, the center of the curve (that is, the 50 percent point) is relatively stable. Therefore, the net effect of the misses of the 0.105 and 0.108 inch cracks is to increase the probability of detection for smaller cracks.

The sensitivity of PoD estimation to "rogue points" is discussed by Hyatt, et al.[20]. They also provide several statistical procedures to downweigh the contributions of extreme points. For the most part, the high influence or possible "rogue" points in the data we gathered were at large crack sizes. The threshold model discussed in section 4.1.1 addresses this issue and is fitted to the data for inspector EC4. The addition of a background miss rate for the lognormal fit resulted in an estimate of 0.01, but the improvement in the fit was not statistically significant. This was not the case for the field inspections that are presented in section 4.2.2. There, the background miss rate substantially improved fits for many of the inspections.

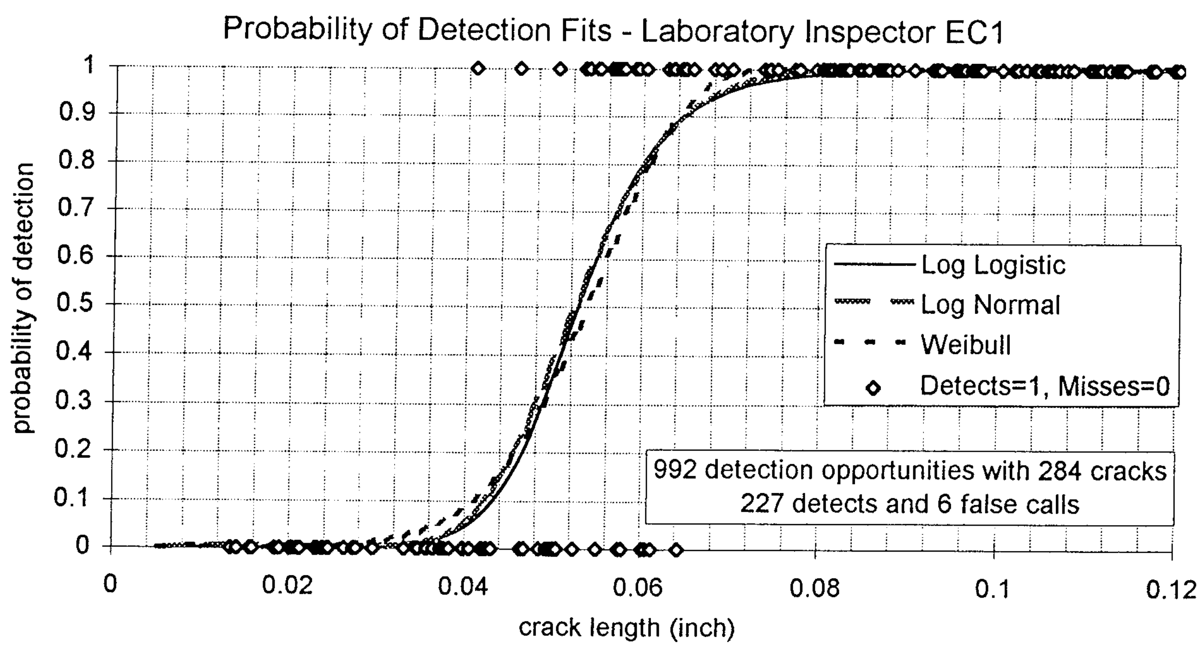

FIGURE 4.1 POD FIT TO LABORATORY INSPECTOR EC1 RESULTS

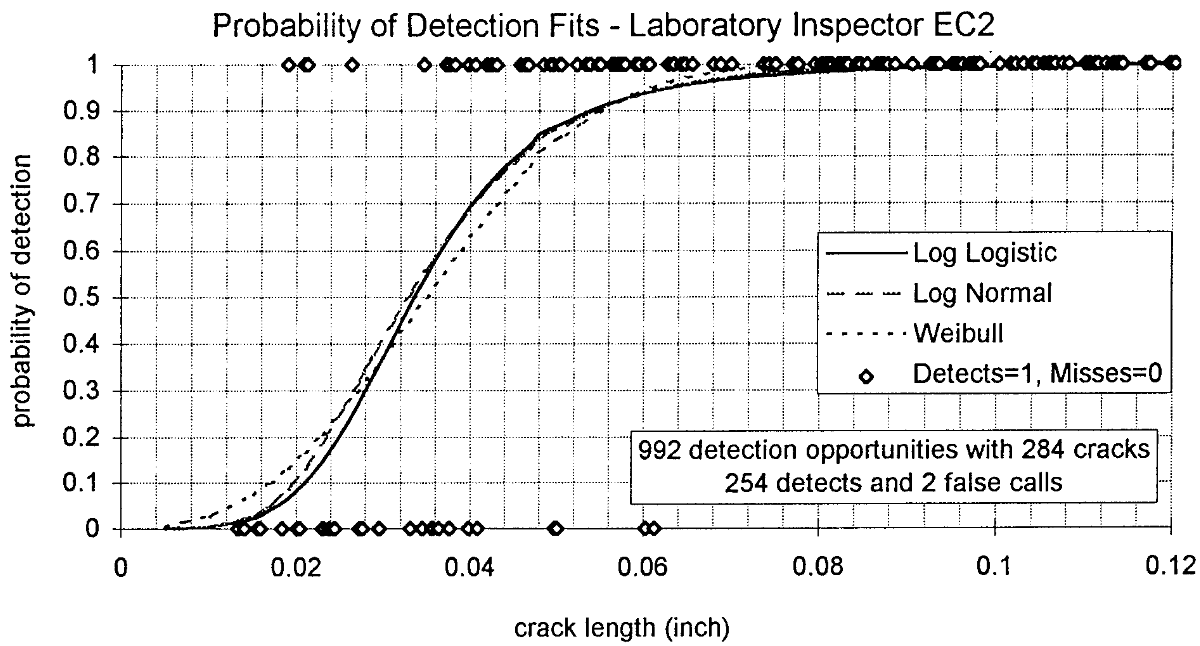

FIGURE 4.2 POD FIT TO LABORATORY INSPECTOR EC2 RESULTS

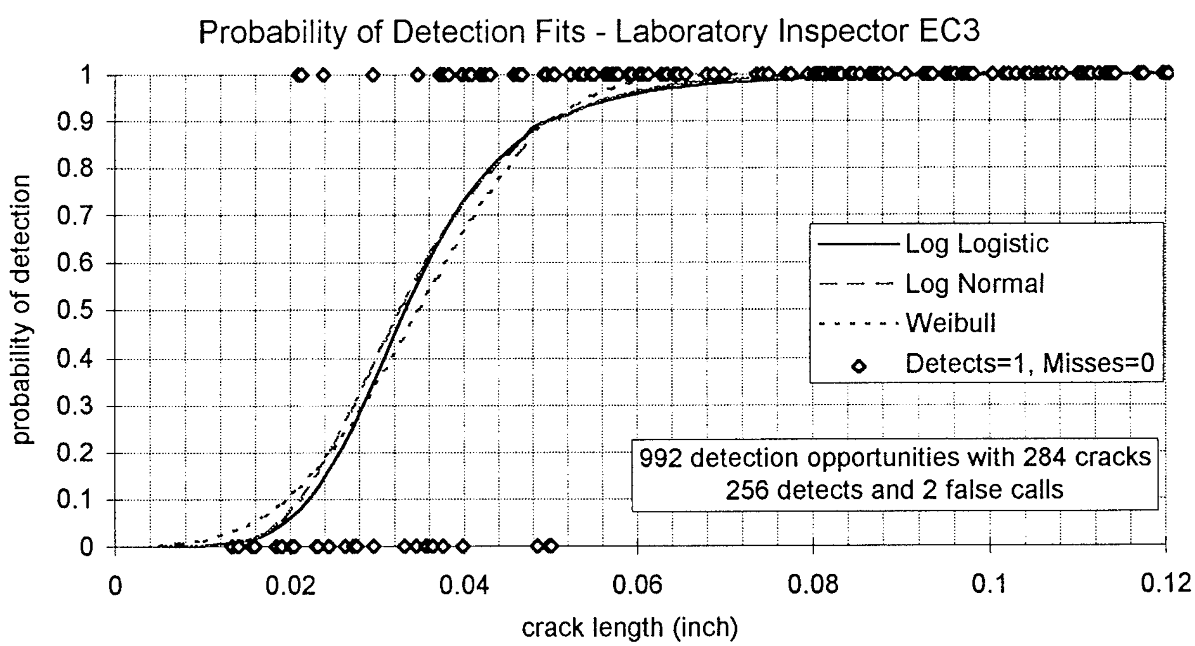

FIGURE 4.3 POD FIT TO LABORATORY INSPECTOR EC3 RESULTS

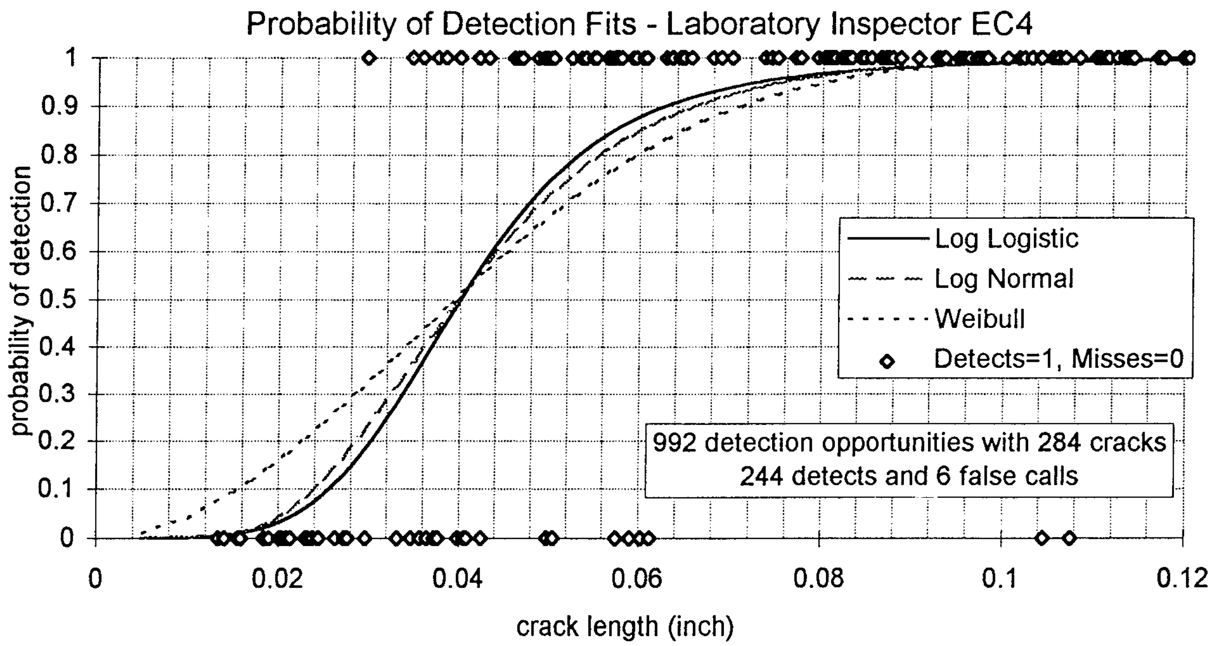

FIGURE 4.4 POD FIT TO LABORATORY INSPECTOR EC4 RESULTS

The similarity between the log logistic and the lognormal PoD fits has been noted by others [10, 11] and has been verified with the data gathered in this experiment. We choose to focus the analysis for the field data by using only the lognormal fits, with and without the background miss rate threshold parameter. Figure 4.5 presents the lognormal fits for the four laboratory inspections in one graph and shows the biggest difference to be in the inspection using the sliding probe.

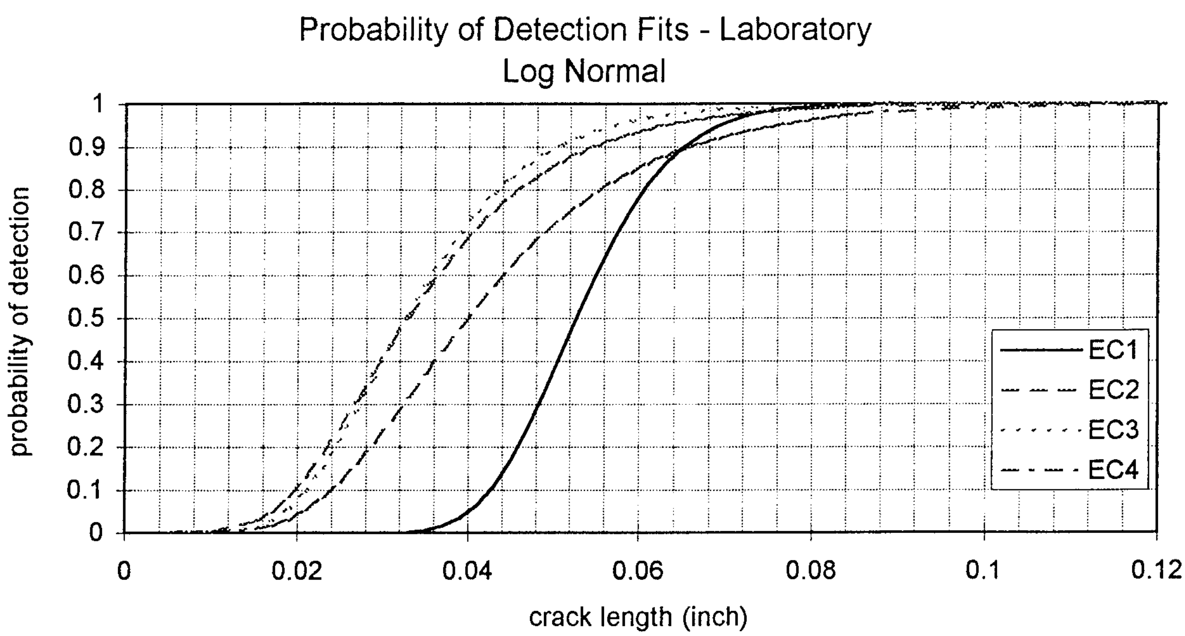

FIGURE 4.5 LABORATORY INSPECTORS POD FITS

Before presenting the field inspection data, we briefly discuss the use of confidence limits in describing PoD curves. Figure 4.6 shows the lognormal fit to EC1 inspection along with an upper 95 percent confidence bound on the fitted probability and a lower 95 percent confidence bound. The method used to calculate these confidence limits is based on the asymptotic normality of the estimates of α and β. The estimated covariance matrix of (α,β) is used to set confidence limits on the quantity $\alpha + \beta \bullet log(a)$, the argument of the normal distribution function [11,19].

In figure 4.6 it is seen that the lower 95 percent confidence bound passes through the probability of 0.90 at approximately 0.072 inch. In this example, 0.072 inch would be the 90 percent probability of detection value at a 95 percent confidence level. This 90/95 crack size has been used to characterize NDE systems. In this example the best estimate of the crack size yielding a probability of detection of 0.90 is 0.065 inch. Thus, the confidence factor adds approximately 0.007 inch to the best estimate. The confidence factor primarily reflects the amount and distribution of cracks used to determine the PoD fits. In the analysis of the field inspection data it will be seen that other factors contribute to PoD curve variations to a much greater extent than the 0.007 inch that is reflected in the individual curve confidence limits. For this reason, confidence limits that primarily reflect the number and size distribution of the cracks will not be given, but rather we will discuss the factors that contribute even more to the overall observed variation.

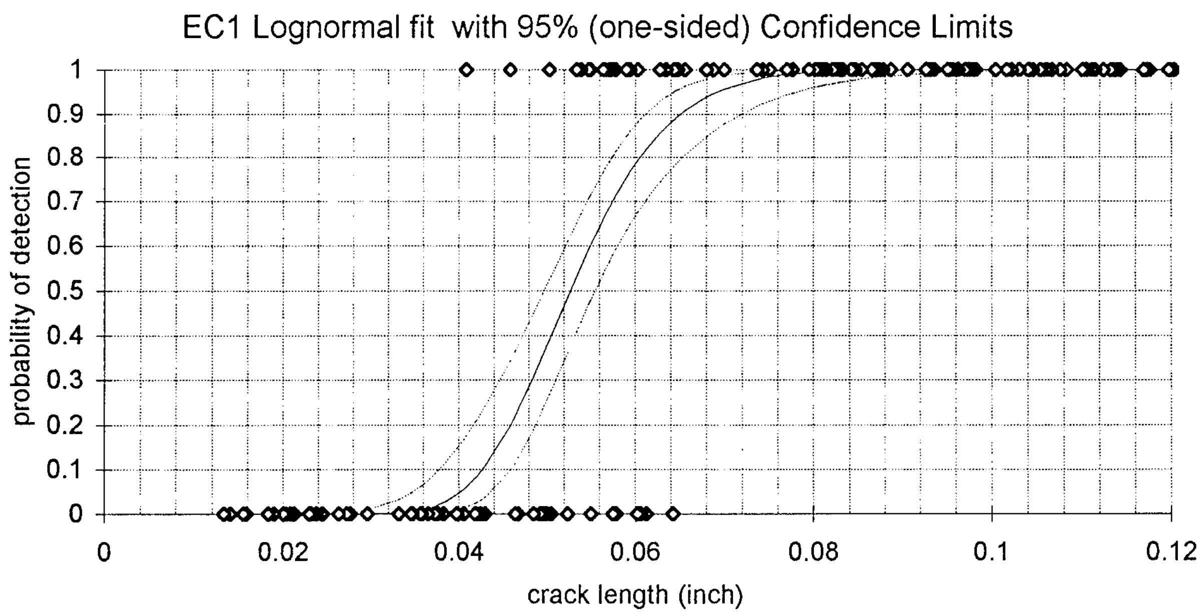

FIGURE 4.6 95 PERCENT CONFIDENCE LIMITS ON INSPECTOR EC1 RESULTS

4.2.2 Field Inspections.

In the pre-inspection briefings, we requested that the inspectors mark the location of cracks around the rivet. The intent was to determine whether all cracks had been detected at a given rivet site. In looking at the data it became clear that many of the rivet sites with cracks from both

sides were only being marked once. Sometimes the rivet site marking was up or down and could not be unambiguously attributed to either crack. Conversations with some of the inspectors verified that they usually did not mark crack orientations.

In general, the inspectors were not (and were not required to be) accurate in specifying crack orientation. Thus, to reflect fairly the detection capabilities, data were analyzed on a rivet site basis rather than by individual cracks. Thus, at rivet sites containing two cracks, the length of the longest crack was used in the analysis. However, the information whether a rivet site had a crack on the right, on the left, or on both sides was retained and included in the overall analysis to determine if there was an effect on the detection probability.

The inspectors were asked to give subjective ratings to their calls. The PoD fits presented here are based on all positive calls regardless of the inspector's rating. (The rating format is discussed in more detail in section 4.3.)

Forty-five inspections were performed at the nine participating facilities. The inspection results are coded in the following manner. The nine facilities are coded A through J (no I). The distinct inspections within a facility are coded 1 through 4 and the repeat inspection at each facility is noted by appending an R to the original code. (Thus, A1R denotes the repeat inspection of A1.) Figures 4.7 through 4.15 present the individual inspection fits by facility.

Large cracks were missed in many of the inspections. (A summary of large cracks missed is given in section 5.) It is likely that many of these misses are the result of factors independent of crack size. For example, momentary distractions or intermittent problems with the inspection equipment could cause a signal to be overlooked. To model this situation, all the PoD curves were fit using the lognormal distribution with a threshold parameter, as discussed in section 4.1.1. The estimated model parameters (with and without the threshold) are given in appendix B. Comparison of fits for several of the inspections are also shown.

The total number of detections and the number of false calls for each inspector are given in the legend of the graphs of figures 4.7 - 4.15. The relationship between false call rates and PoD is discussed in more detail in section 4.3 The "lighter" curves in figures 4.7 - 4.15 are for the set of inspections done by the same inspector(s).

The equipment and procedures used by the inspectors are given in table 4.2. (Specific associations with the inspections are not given inorder to preserve the confidentiality promised to the inspectors.) It is apparent that a wide range of equipment and calibration standards were employed.

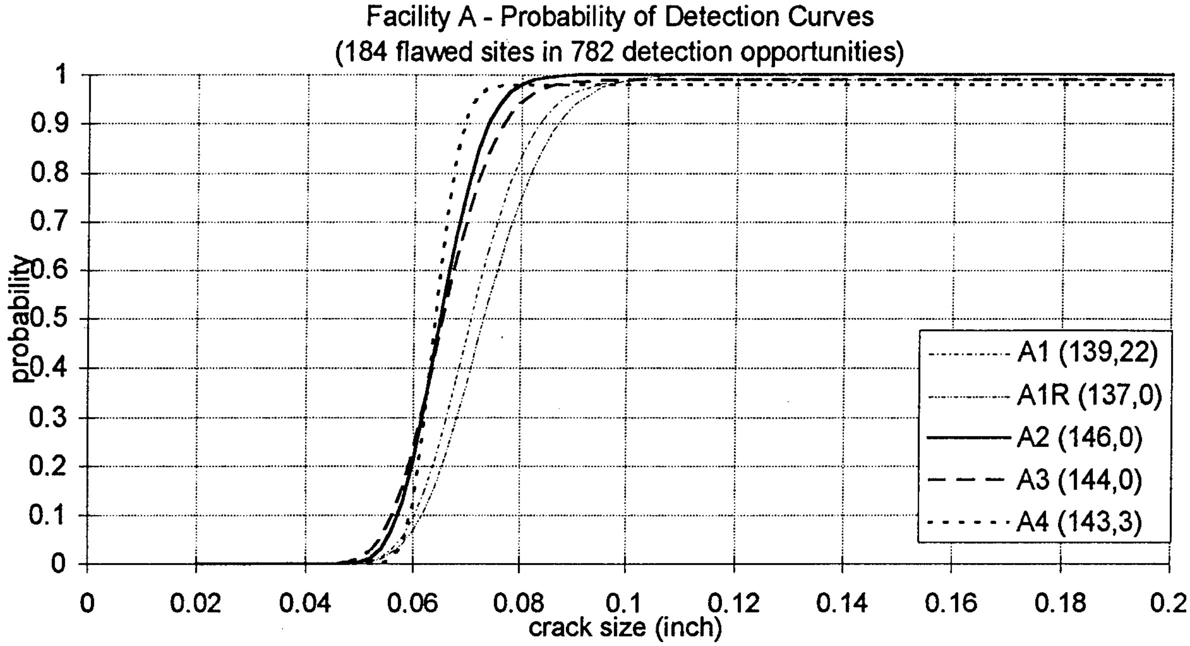

FIGURE 4.7 FACILITY A INSPECTIONS WITH (# DETECTS, FALSE CALLS)

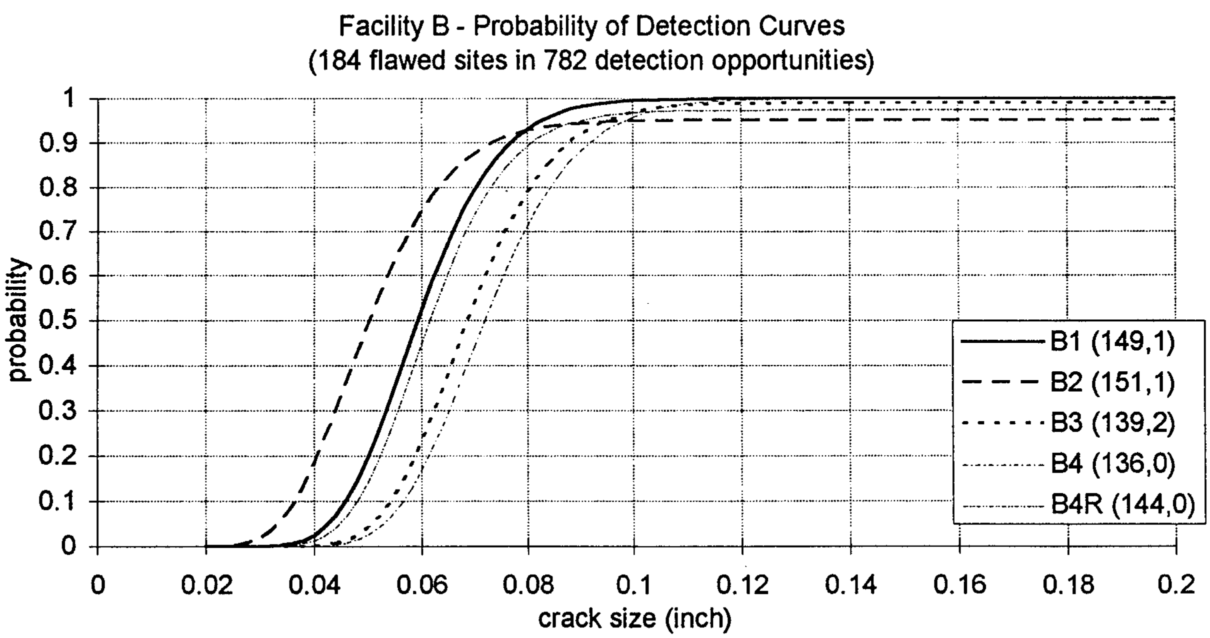

FIGURE 4.8 FACILITY B INSPECTIONS WITH (# DETECTS, FALSE CALLS)

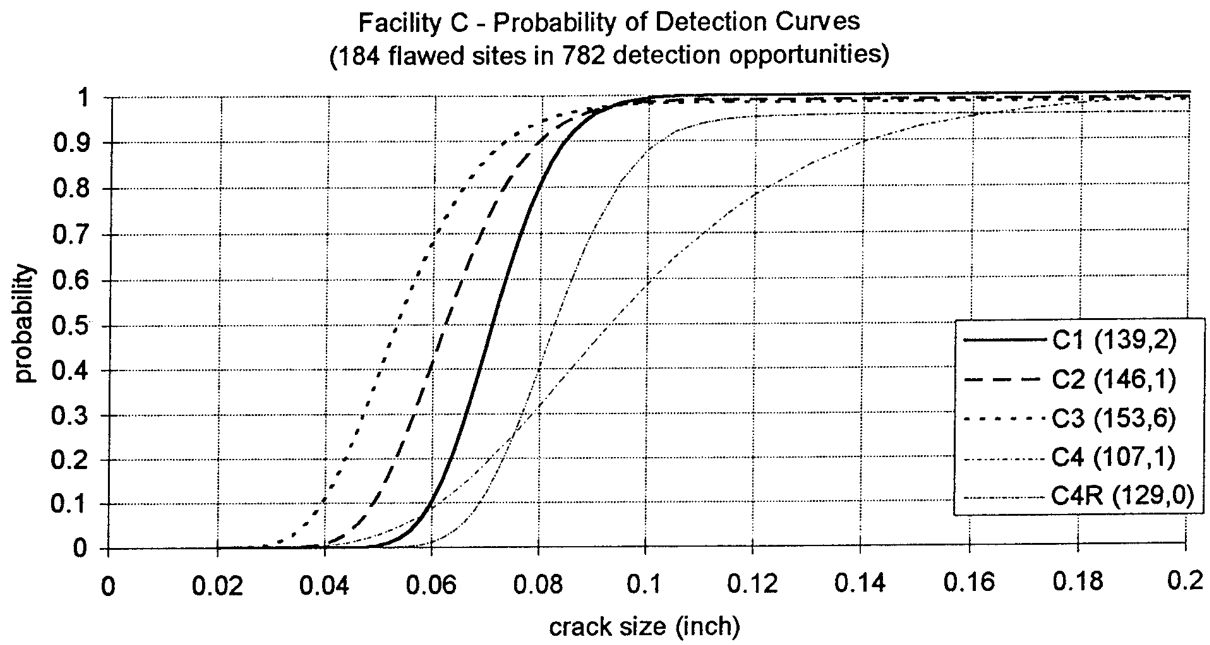

FIGURE 4.9 FACILITY C INSPECTIONS WITH (# DETECTS, FALSE CALLS)

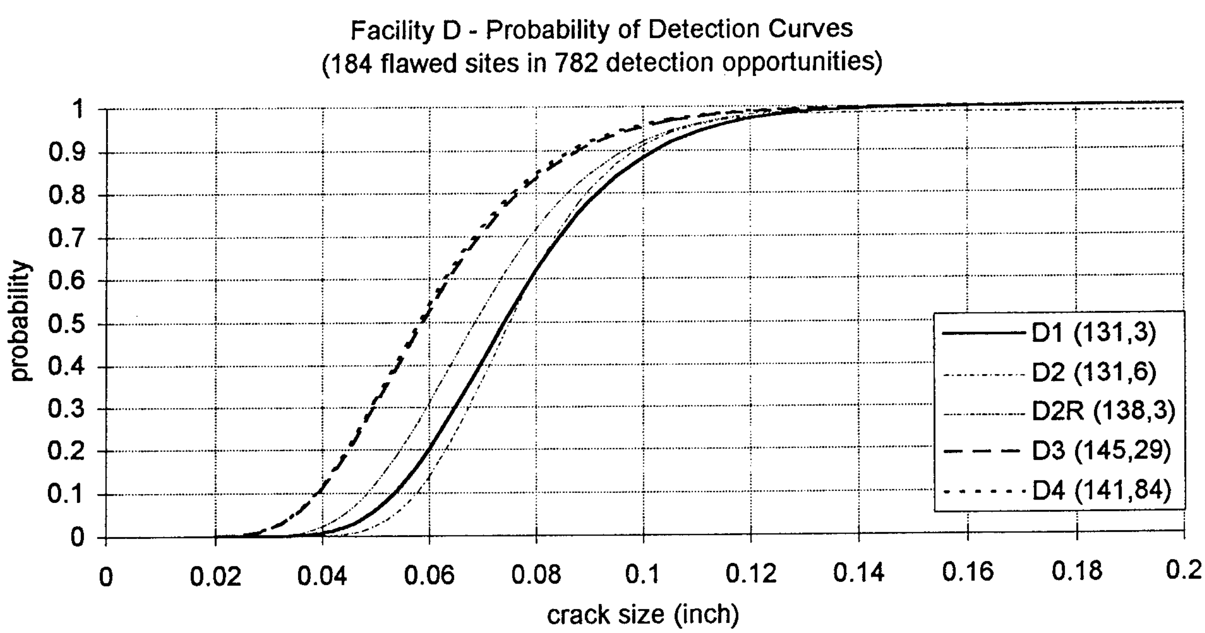

FIGURE 4.10 FACILITY D INSPECTIONS WITH (# DETECTS, FALSE CALLS)

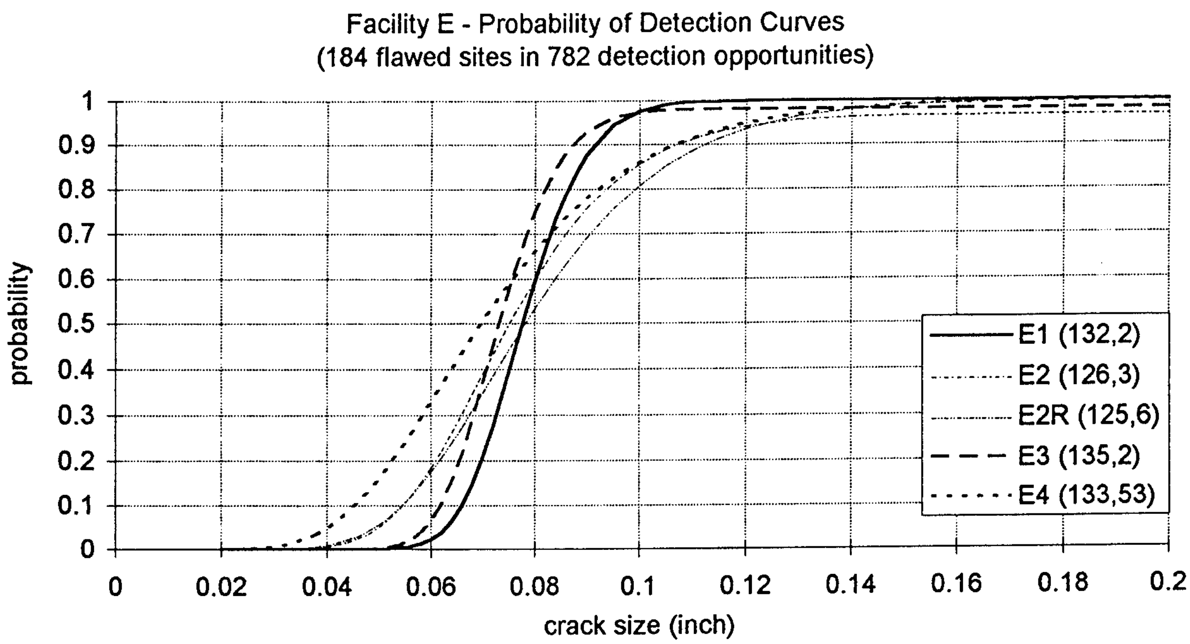

FIGURE 4.11 FACILITY E INSPECTIONS WITH (# DETECTS, FALSE CALLS)

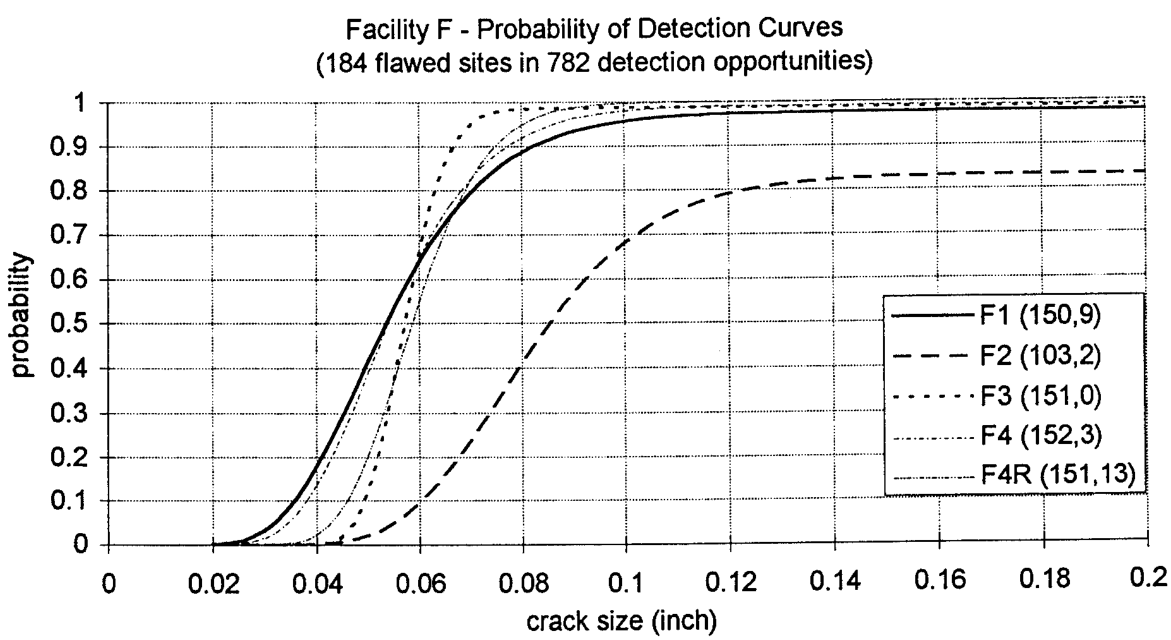

FIGURE 4.12 FACILITY F INSPECTIONS WITH (# DETECTS, FALSE CALLS)

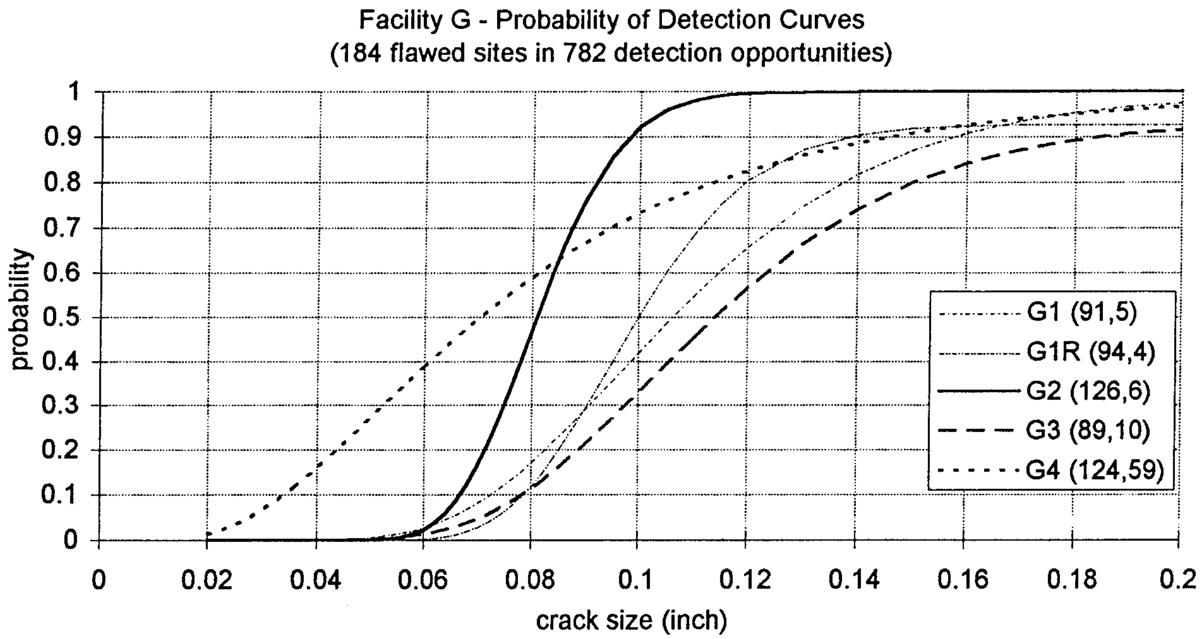

FIGURE 4.13 FACILITY G INSPECTIONS WITH (# DETECTS, FALSE CALLS)

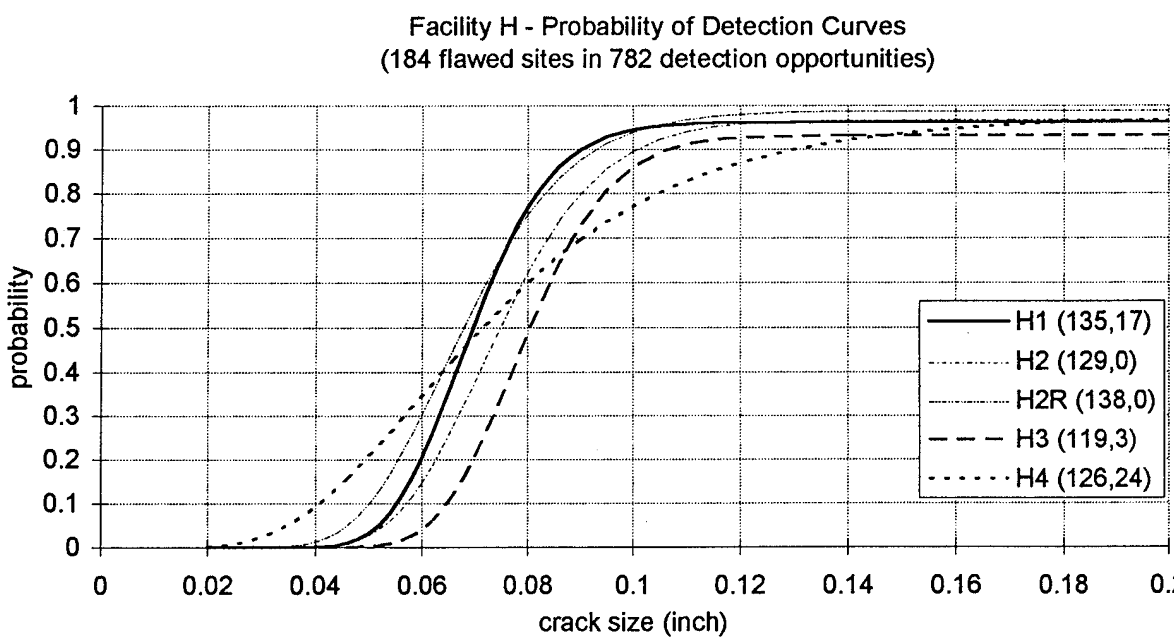

FIGURE 4.14 FACILITY H INSPECTIONS WITH (# DETECTS, FALSE CALLS)

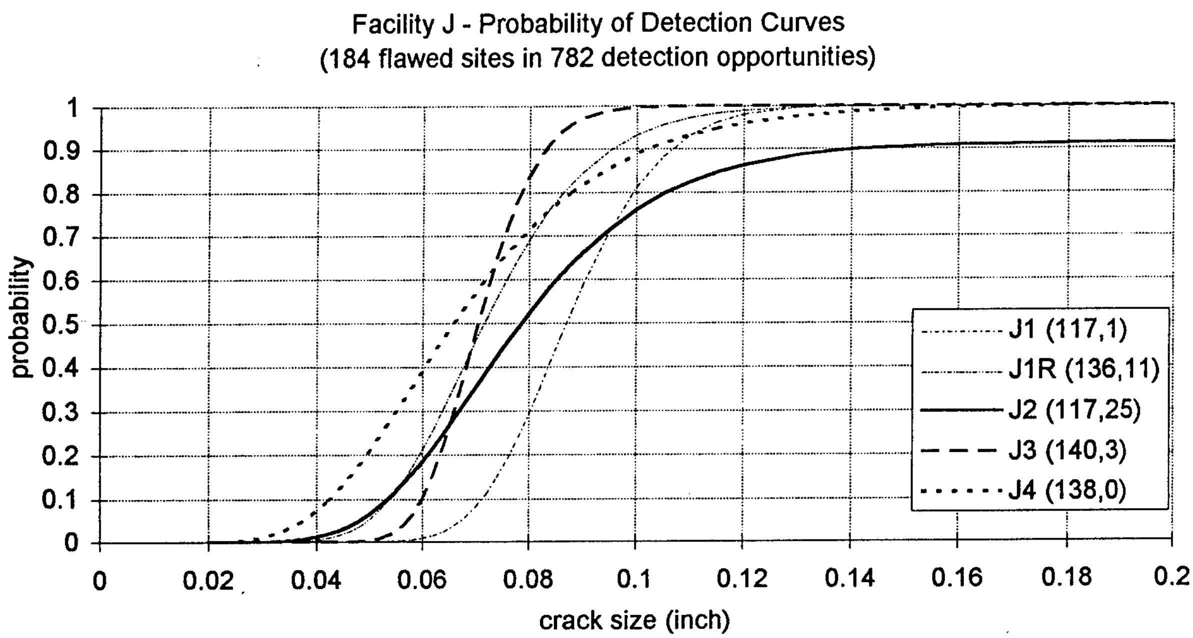

FIGURE 4.15 FACILITY J INSPECTIONS WITH (# DETECTS, FALSE CALLS)

The experiment monitors traveled with an eddy current tester and with a "master" #369 calibration standard. The calibration standard(s) being used at each of the facilities were compared to the "master." The intent was to characterize the contribution of the variation in the standards. When each inspector within a facility uses the same standard, the inspector-to-inspector variation observed in that facility would not be due to reference standard variations. Given the extent of between inspector variation and the qualitative differences in calibrations between facilities (e.g., the use of notches on universal blocks), the variation between the few characterized reference standards is only a small part of the total variation.

Comparisons of the fits in figures 4.7 through 4.15 indicate that Facility G had the overall poorest performance. (This is borne out in a more formal analysis.) Facility G was one of the three facilities where the inspectors did not calibrate their equipment using the calibration blocks called for in the Boeing procedures [9]. The inspectors set up on a notch on a universal eddy current standard rather than the Boeing designated standard that replicated the rivet site structure. Facility A also followed similar procedures (however, using a different size notch) but had one of the overall best performances. Of course, other procedural differences existed between the facilities. For example, both employed the oversize template procedure, but the inspectors at Facility A used a 15/32-inch template, while the inspectors at Facility G used a 9/16-inch template. The same equipment was used by the inspectors within a facility, but different equipment was used at each facility. Within the context of this experiment such observed differences can only suggest factors with potential influence. More formalized study would be necessary to separate and establish the effects of these uncontrolled observed factors.

Table 4.2 *Equipment and methods used in inspections (random order)*

[1]Method	[2]EC Equip.	[3]Probe Reference #	Freq. (kHz)	[4]Calibration standard	Template size	[2,3]Verification Equip. + Probe Reference	# occur
T	N19	MP901-50B	200	NORTEC TB-SI	15/32"		3
T	N19	MP457-50C	200	NORTEC TB-SI	15/32"		
T	N19	MP457-50C	200	#369	15/32"		
S	N19	SPO-3806	30	#290	7/16"	MIZ10	
S	N19	SPO-3806	28	#290	7/16"	N19 + VM229	
S	N19	SPO-3806	32	#290	7/16"	N19 + VM229	
S	N19	SPO-3806	25	#369 , #290	freehand	N19 + VM229	
S	N19	SPO-3806	30	#290	freehand	N19 + VM229	
S	N19	SPO-3806	38	727 ET 108AF	freehand	ED520 + MP905-50	2
S	N19	SPO-3806	38	#290		ED520	
S	N19	SPO-3806	38	727 ET 108AF	freehand	ED520	2
T, S	N19	SPO-3806	30	#369	15/32" & 1/2"	N19 + MP902-40B	
T	L UH	MP905-50B	200	#369	15/32" & 1/2"		2
S	N19	SPO-3806	27	#369	1/2"	N19 + MP455-50C	
T	L UH	MP905B	200	#369	15/32" & 7/16"		
T	ED520	MP902-505X	[5]FIXED	On panel	7/16"		5
S	MIZ22	LTW 1004-2	26	#290 Ptd & Unptd	7/16" & 1/2"	no verification	
S	MIZ22	LTW 1004-2	26	#290 Ptd & Unptd	1/2"	ED520 + MP455-50C	
S	MIZ22	LTW 1004-2	26	#290 Ptd & Unptd	7/16" & 1/2"	MIZ22 + MP905-50	
S	MIZ22	LTW 1004-2	26	#290 Ptd & Unptd	7/16", 1/2" & freehand	ED520 + MP455-40C	2
T	MIZ20	VM200F	210	VM PNSB095300165	9/16"		5
S,T	MIZ20	SPO-2210	24	#290	7/16"	MIZ20 + GK1RR90F6	
S+T	MIZ20	SPO-2210	24	#290		For2.8 + FCU-RAF .5x6	
S	MIZ20	SPO-2210	24	#290	7/16"	For2.8 + FCU-RAF .5x6	2
S	MIZ20	SPO-2210	24	#290	7/16"	MIZ20 + NDT probe	
T	N19	MP902-40B	250	#369	13 mm		
T	N19	MP905-50B	250	#369	12 & 13 mm		
T	MIZ20	MP905-50B	260	#369	7/16"		
T	MIZ20	MP905-50B	260	#369	13 mm		
R	E B1	22 SSRO	460	#369	Rotation Dia. 0.325"		

Table Notes:

1. **T** - Initial inspection done using template procedures. **S** - Initial inspection done with sliding probe. **R** - Rotating surface probe. Combination **S, T** denote inspections where the sliding probe was abandoned during the inspection, **S+T** all sited checked with both procedures.
2. Equipment: **N19** - Staveley Nortec 19, **L UH** -Hocking Locator UH, **ED520** - Magnaflux ED-520, **MIZ22** - Zetec Miz-22, **MIZ20** - Zetec Miz-20, **E B1** - Rohmann Elotest B1, **For2.8** - Forester 2.8.
3. Probe manufacturers: **MP...**- NDT Product Engineering, **SPO....**- Nortec, **LTW....**- NDT Eng. Corp., **VM....**- VM Products, Inc., **FCU...**-Tyvin Probe
4. Calibration blocks: **#369** - Boeing I.D. for two sheets of 0.040-inch-thick 2024-T3 or T4 Al Clad material fastened with BACR15CE5D or BACC15CE6D fasteners. Contains one row of 5/32-inch fasteners and one row of 3/16-inch fasteners. An edm notch 0.007 inch wide and 0.100 inch is present in each row.
   #290 - Boeing I.D. for reference standard similar to #369, but also containing cracks at 60 degree angles off-horizontal. Nortec and VM are universal eddy current standards.
5. ED 520 does not have adjustable frequency. Frequency operates in range of 70 to 150 KHz.

4.3 ROC INSPECTION RESULTS.

As was discussed in section 4.1.3 an ROC analysis addresses the false call rates as well as detection rates. False call rates were given in the PoD curve legends of figures 4.7 - 4.15. Here, we present false call data in a more formal manner.

For an ROC analysis, different criteria or decision-making levels for calling crack indications are needed. These were obtained by asking each inspector (or inspection team) to say how "confident" they were that the eddy current indications should be reported. A three-point scale was asked for, with the following guidelines being given (See RAE 9 of appendix A):

3 means you are absolutely certain that the indication is reportable.
2 means that you are reasonably sure that the call of an indication is correct.
1 means that you have some doubts about the indication being reportable, but that you cannot overlook it.

Three points on an ROC curve for each inspection are determined in the following manner. The first point considers a call as being made only if it was rated a 3. The proportion of detections is then plotted against the proportion of false calls among the 3's. The second point is determined by including the 2's with the 3's and with this expanded set determining the proportion of detections versus the proportion of false calls. The third point is determined in a similar manner, but including all the calls.

The three points reflect a changing decision criterion. In deriving PoD curves any one of the three criterion levels could be used, thereby obtaining a family of PoD curves. Individual PoD curves derived from the loosest criterion (3-2-1 group) were given in section 4.2. This represents the basic capability of the inspection system. Inspectors may have had doubts about the validity of some calls, but they believed that their inspections were providing some evidence of cracks. All the PoD curves of section 4.2 were derived from this group.

In figure 4.16, a scatterplot of the overall detection rate versus the overall false call rate for the 45 inspections is shown. Three inspections (D4, G4, and E4) have false call rates that stand apart from the rest.

The subjective ratings of 1, 2, or 3 provide further information concerning the false calls and their relationship to true calls. However, the use of subjective ratings was not uniform across the inspections. In ten (or 22 percent) of the inspections, no use was made of the subjective scale. In another 24 (or 53 percent) of the inspections fewer than 10 percent of the calls were rated 2 or 1. (A breakdown of the number of calls and false calls by ratings is given in table B.2 in appendix B.)

ROC curves for eight inspections are shown in figures 4.17 to 4.24. Three curves are shown in each figure. The curves are for the cracks with lengths less than 0.050 inch (still under the countersunk rivet head for the most part), cracks with lengths between 0.050 and 0.100 inch, and cracks that exceed 0.100 inch in length. The 0.100 inch level was chosen because Boeing procedures call for setting up the inspection equipment using standards with 0.100 inch cracks. Because the procedures call for setup to a standard of length 0.100 inch, cracks of this length or

greater should have a high probability of being detected. The curves include the (0,0) and (1,1) points as natural endpoints to extreme criterion levels, although none of the inspections were performed at these extremes.

The ROC curves shown are for 8 of the 11 inspections where 10 percent or more of the calls were rated 2 or 1. The three inspections in this category that are not shown (C4R, F2, and J4) had no or very few false calls.

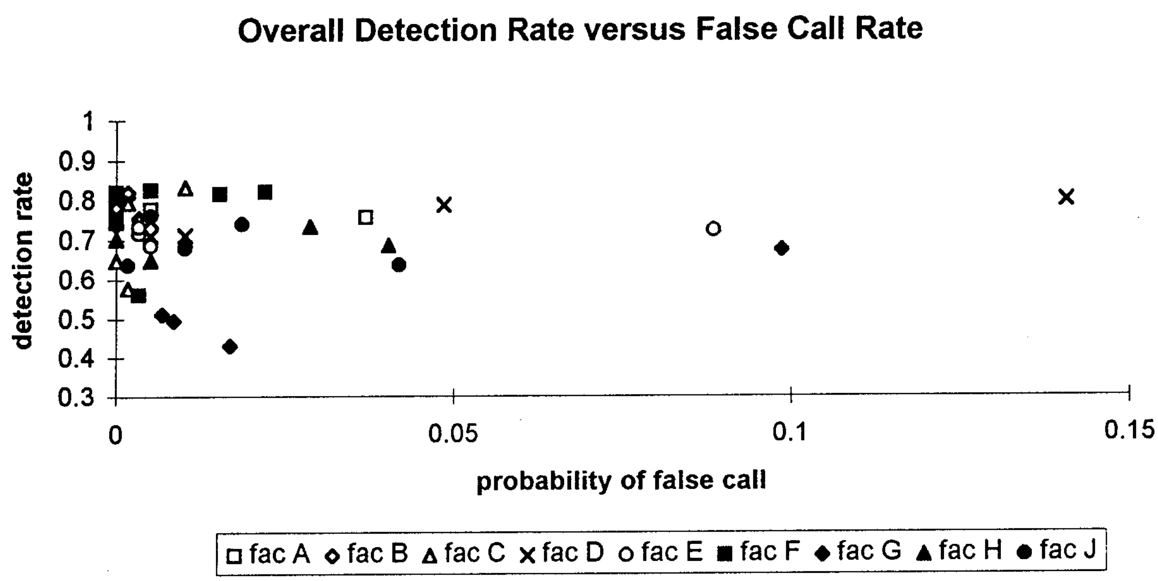

FIGURE 4.16 DETECTION RATE VERSUS FALSE CALL RATE FOR (3,2,1) CALLS

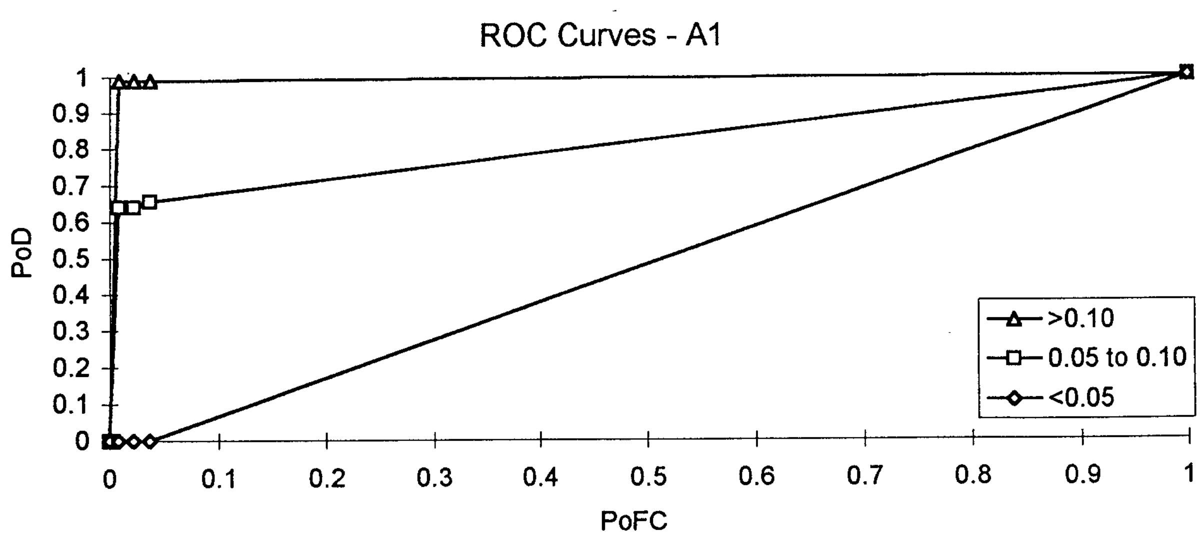

FIGURE 4.17 ROC CURVES FOR INSPECTION A1

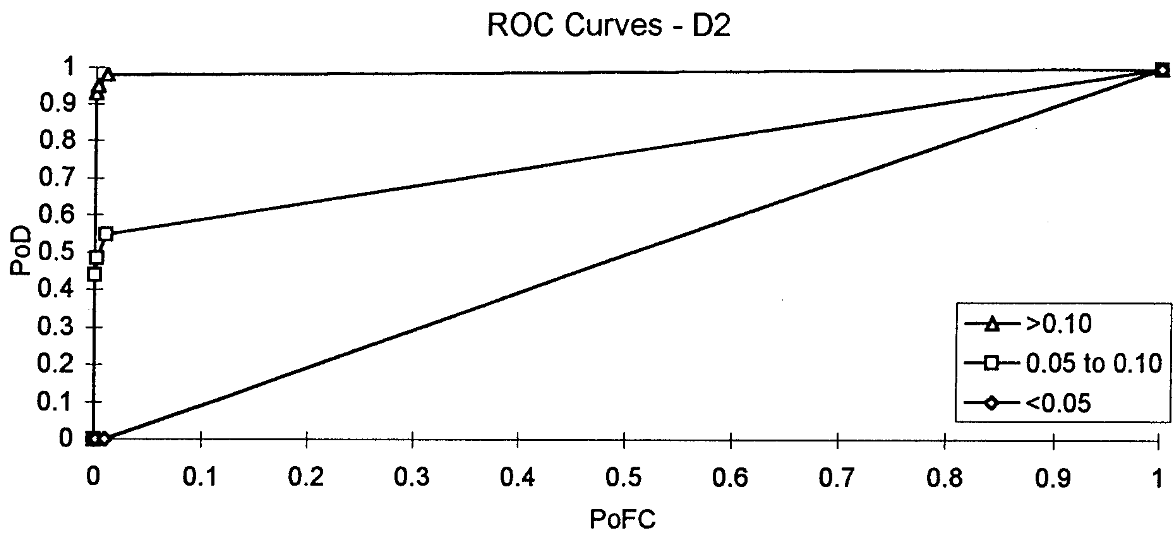

FIGURE 4.18 ROC CURVES FOR INSPECTION D2

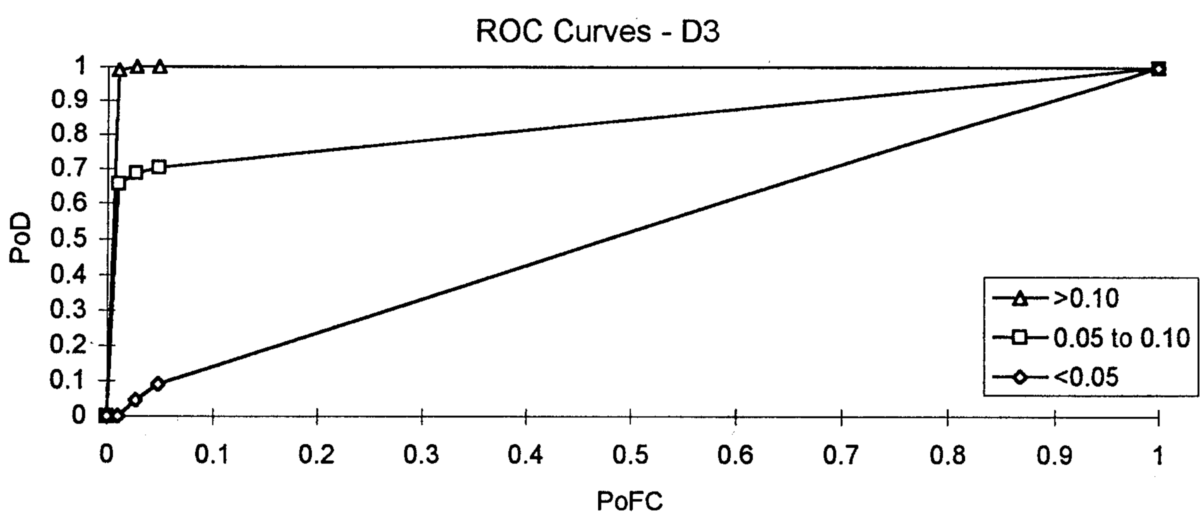

FIGURE 4.19 ROC CURVES FOR INSPECTION D3

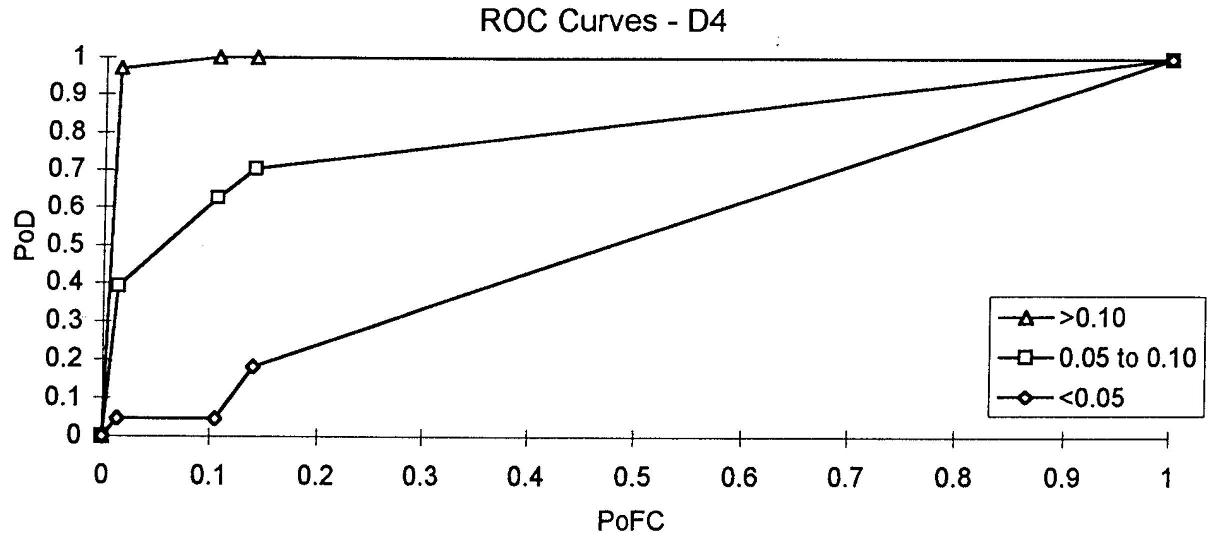

FIGURE 4.20 ROC CURVES FOR INSPECTION D4

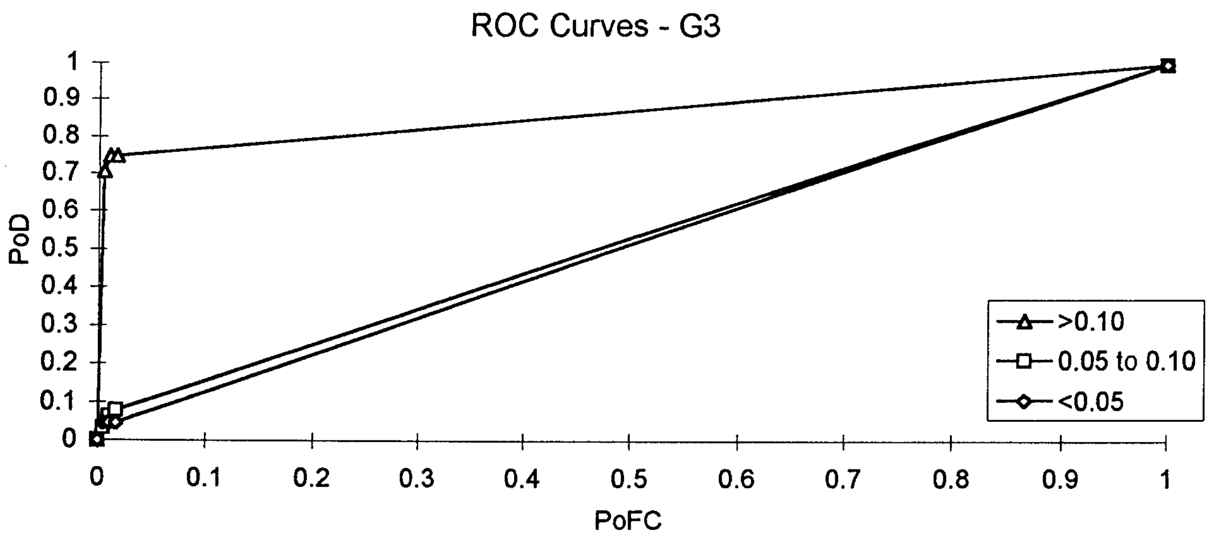

FIGURE 4.21 ROC CURVES FOR INSPECTION G3

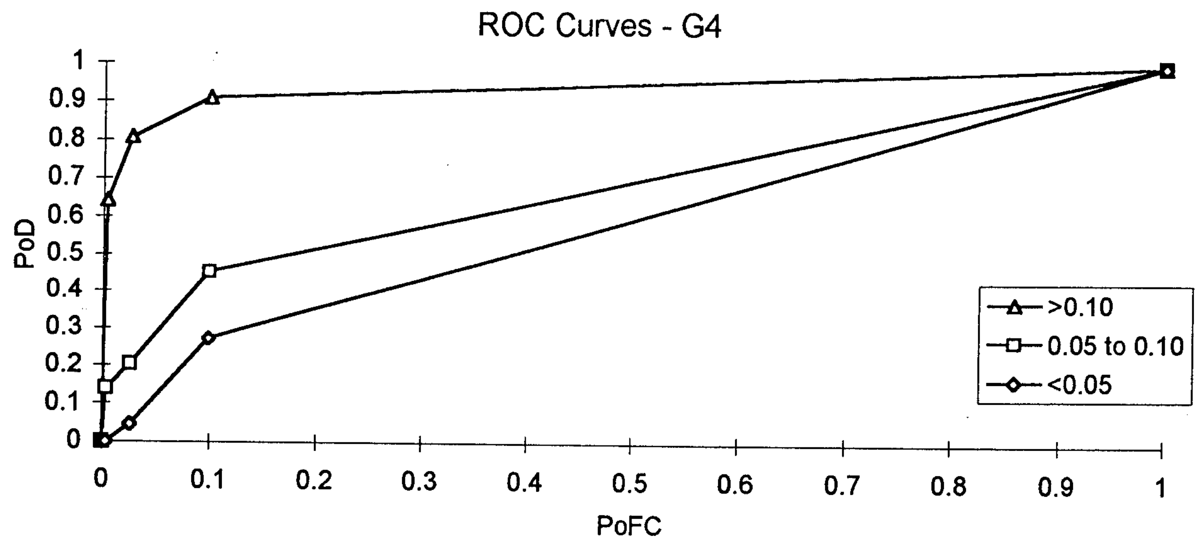

FIGURE 4.22 ROC CURVES FOR INSPECTION G4

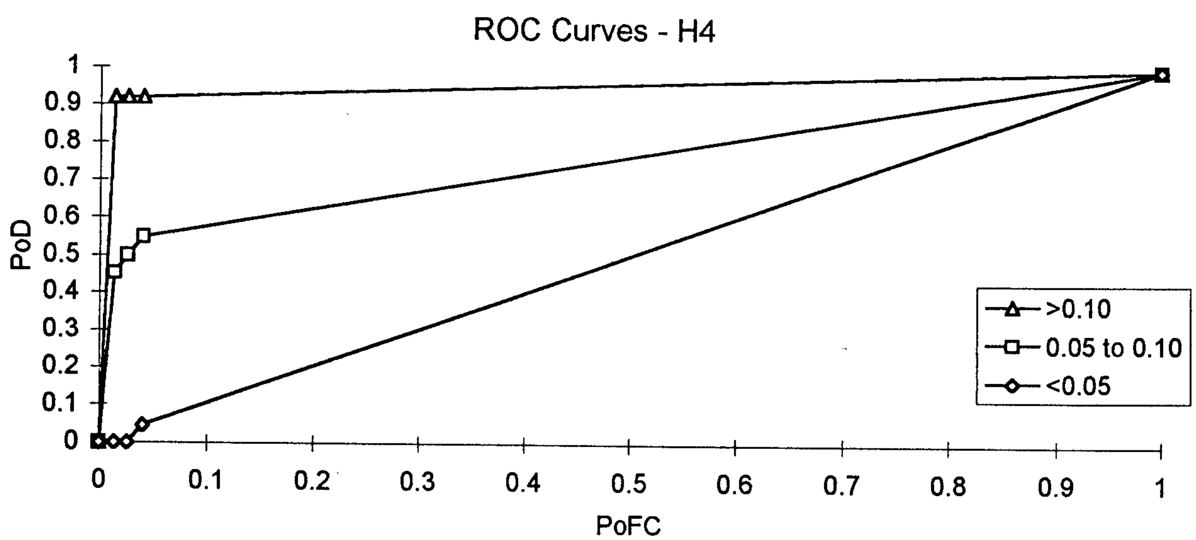

FIGURE 4.23 ROC CURVES FOR INSPECTION H4

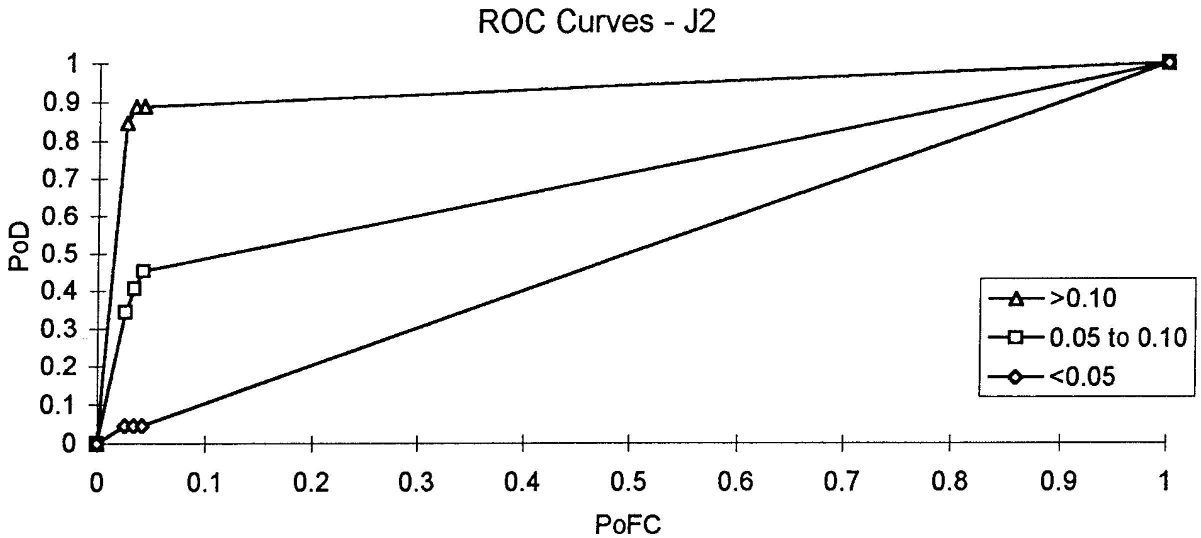

FIGURE 4.24 ROC CURVES FOR INSPECTION J2

There is information in the ROC plots about criteria shifts of the inspectors. For example, in lowering his criterion (going from 3's to 3's and 2's and finally to 3's, 2's, and 1's), inspector A1 (figure 4.17) primarily added false calls. That is, the inspector did not detect many additional cracks. The implication is that the PoD curve for inspector A1 given in section 4.2.2 (based on 3's, 2's, and 1's) could be maintained with a decrease in the overall false call rate. This was borne out by A1's repeat inspection where no use was made of the 2 or 1 ratings and the false call rate was 0, with only a slight shift in the PoD estimated curve.

The same point can be made about inspector D3 concerning the larger cracks. However, Inspector D3 picked up some of the smaller cracks in about the same proportion as false calls were being made. The effect of the more strict criterion in D3's case would be to shift the estimated PoD curve to the right with a greater shifting in the lower tail (i.e., where PoD < 0.5)

If inspectors D4 and G4 applied a stricter criterion in making positive calls, substantial shifts in their PoD curves would occur. Thus, substantial detections at the lax criterion were gained by adding false calls. Tightening the criterion would result in PoD curves more in line with the other inspectors in their respective facilities, with fewer false calls.

One of the extreme inspections regarding false calls was that of E4. Inspector E4 did not use the subjective ratings of 2 and 1. Normally this would mean that no information was available about the relationship between PoD and PoFC. However, the experiment monitors observed the inspector making an internal criterion shift due to an overheard conversation. The NDT manager for the company came onto the hangar floor while the inspection was proceeding. Within hearing distance of the inspector, he had a conversation with another person concerning the distribution of crack lengths that should be capable of being detected. The experiment monitor noted a distinct shift in the criterion the inspector used in making a call. Before hearing the conversation, E4 had correctly marked 9 of 19 cracks in the 0.050 to 0.10 inch range. He had made no false calls in 304 unflawed sites. After overhearing the conversation, the inspector correctly marked 16 of 25

cracks in the 0.050 to 0.10 inch size range, but also made 53 false calls in 294 unflawed sites. Figure 4.25 shows the resultant ROC curve for this crack range.

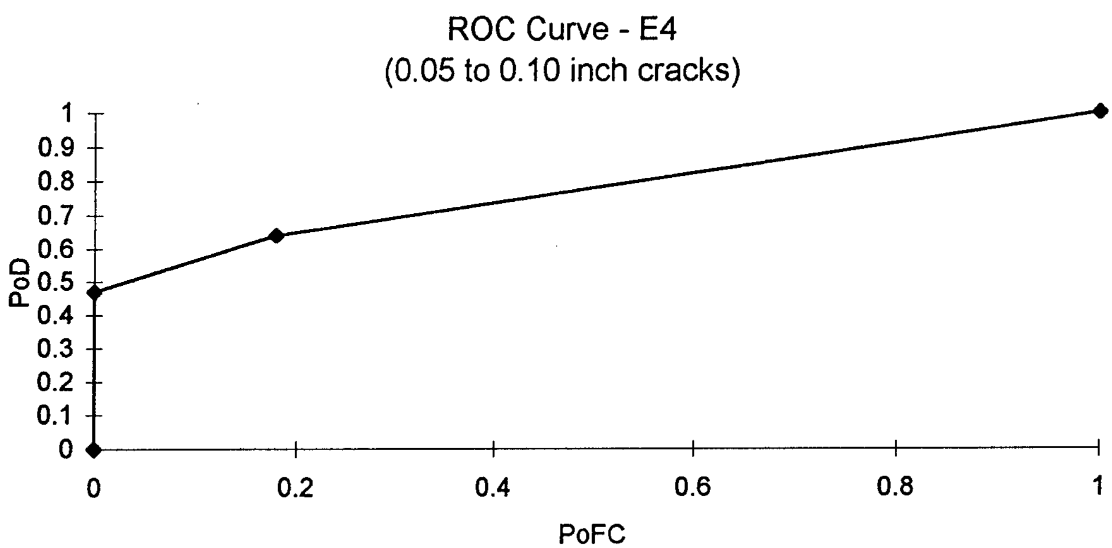

FIGURE 4.25 E4 ROC CURVE FROM OBSERVED IMPLICIT CRITERION SHIFT

Inspector E4 presents an extreme case of a general phenomenon that can occur during any inspection. The phenomenon is a variable detection criterion being applied over time. This is more likely to occur with those inspectors that do not establish explicit detection criterion (by gating alarms or drawing decision lines directly on instrument displays) during setup.

4.4 ANALYSIS FOR DESIGN FACTOR INFLUENCES.

In this section we consider the effects of the various factors discussed in section 2.2. Before giving a full analysis, we look at detection curves fit to the composite data. In figure 4.26 the proportion of time that each of the 184 flawed sites was detected is shown versus the length of the crack. Also shown are the lognormal curves that are fitted to the data with and without the threshold parameter. It is stressed that the PoD so represented is an average across inspectors and conditions. Each of the different inspections has a different PoD as was shown earlier.

From figure 4.26 it is seen that the shape of the lognormal distribution function captures the shape of the general increase in probability of detection as a function of crack length. There is still substantial variation in detection levels for cracks of nominally the same length. The overall effect of factors discussed in section 2.2 are considered in more detail in a sequential regression analysis.

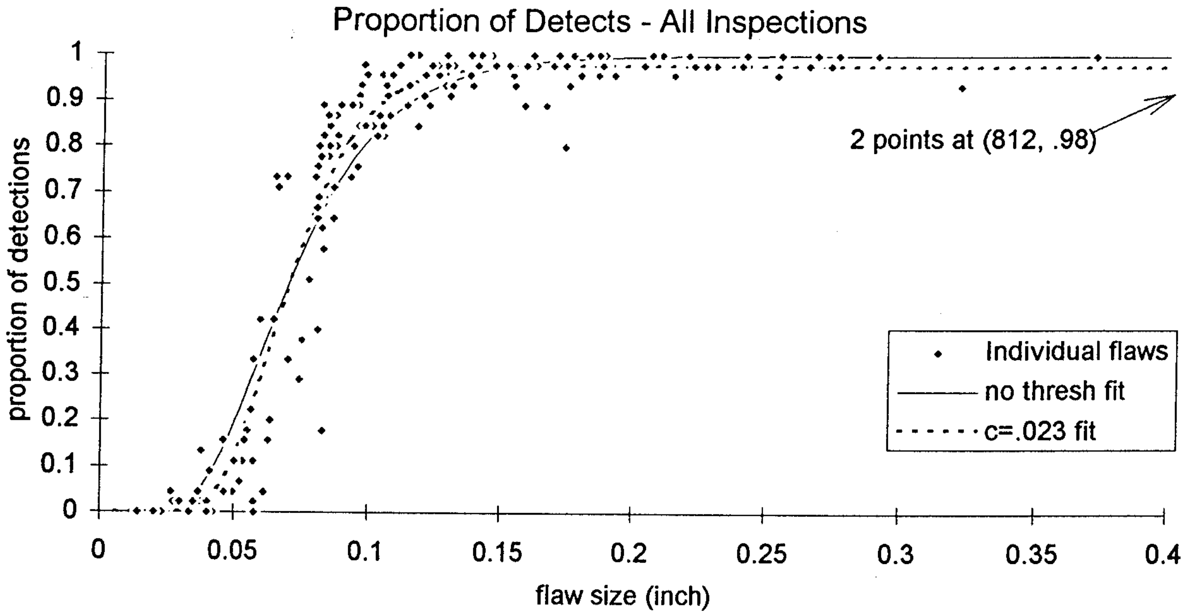

FIGURE 4.26 PROPORTION OF DETECTS FOR ALL INSPECTIONS - WITH LOGNORMAL POD FITS

4.4.1 Sequential Regressions.

We extend the PoD model used for the single inspections to include parameters representing the various experimental factors discussed in section 2.2. The model fit is given by:

$$\mathrm{PoD}(a) = (1\text{-}C)\bullet\Phi(\,\alpha + \Sigma\alpha_i\bullet \mathbf{I}_i + (\beta + \Sigma\beta_i\bullet \mathbf{I}_i)\bullet log(a)), \qquad (3)$$

where Φ is the standard normal distribution function and the $\mathbf{I}_i$ denotes an indicator variable (i.e., $\mathbf{I}_i$= 1 or 0 according to the presence or absence of a specific factor level). A model containing only terms of the form $\Sigma\alpha_i\bullet \mathbf{I}_i$ along with a single β coefficient would allow for shifts in the PoD curve, but not changes in the basic shape of the curve. The $(\Sigma\beta_i\bullet \mathbf{I}_i)$ term allows for the PoD curve to sharpen or flatten with the inclusion of each variable. The additional β terms can be thought of as "slope" interactions.

Besides the factors listed in section 2.2, the inspection procedure initially followed (rotating probe, sliding probe, or template), and whether the inspection was done by a team were added as potential explanatory factors. (NOTE: For the most part, the same factor level applies to all inspections of a given facility for these two parameters.) Parameters that account for the factors are considered in the model and checked for significance. Factors that are not significant are dropped and the remaining factors are fit again. This process is continued until a set of factors is determined that are all judged to be significant. Details of the process are given in appendix B.

Here we give the final model and discuss the implications of the model. The final model depends on the following factors: crack angle, inspection surface/facility (painted or bare), accessibility (or

lap splice position), number of flaws at the inspection site, procedure/facility (sliding probe or template), and the crack density. The facility is included in the inspection surface and the procedure factors to emphasize that the inspections at each facility were all performed on the same surface condition for the small test specimens. Also, for the most part, the inspections within a facility were all carried out using the same procedure. (Further discussion of facility influence on the surface and procedure effects is given in section 4.4.2.)

Each of the noted factors has two levels as given in table 4.3. The indicator variable given in the table takes on the value of 1 if the factor is at level 1 and 0 if the factor is at level 0.

***Table 4.3** Model Factor Levels*

	crack **angle** (I_{ang})	**surf**ace condition (I_{surf})	lap splice **pos**ition (I_{pos})	**num**ber of cracks at rivet (I_{num})	**proc**edure followed (I_{proc})	**dens**ity of cracked rivet sites (I_{dens})
level 0	11 to 22 degs.	painted	low	2	template	low (~10%)
level 1	horizontal	bare	high	1	sliding	high (~40%)

By using the indicator variables defined in the table, the final equation is given by:

$$PoD(a) = (1-.024)\bullet\Phi(^{-}12.0751 + 0.4313\cdot I_{ang} + {}^{-}1.5227\cdot I_{surf} + 0.1637\cdot I_{pos} + {}^{-}0.4546\cdot I_{num} + 0.6622\cdot I_{proc} + {}^{-}3.3598\cdot I_{dens} + {}^{-}.6634\cdot I_{surf}\cdot I_{proc} + {}^{-}.4392\cdot I_{surf}\cdot I_{num} + 0.3660\cdot I_{num}\cdot I_{dens} + (2.7532 + 0.6233\cdot I_{surf} + 0.7049\cdot I_{dens})\cdot log(a)),$$

where Φ is the standard normal distribution function and the crack length, a, is given in mils (thousandths of an inch). For example, the term 0.3660 would be included only if the number of cracks was one (I_{num} =1) and crack density is high (I_{dens}=1).

Note that the value of each of the factors has to be given to specify the model completely. This makes the answer to "What is the effect of the accessibility or position factor?" depend on other factors. However, we answer this question by looking at the effect averaged across the other conditions. There are six factors at two levels each. Therefore, there are 64 distinct combinations of the different factor levels. We set I_{pos}= 0 (i.e., lower row) and calculate the a_{50} and the a_{90} values for each of the 32 other combinations for the remaining factor levels. These values are averaged and are given in table 4.4. The same procedure is followed for I_{pos}= 1 (upper row). From table 4.4 it is seen that the average effect of the lower row (as compared to the upper row) was to add about 5 mils (0.005 inch) to the length of crack detected at the 0.90 rate. This procedure is followed for each of the individual factors to produce table 4.4.

In the analysis the density of cracks was significant, but from table 4.4 it is seen that the average of this effect is small. Note that the 50$^{\underline{th}}$ percentile for "high" density is 0.003 inch larger than the 50$^{\underline{th}}$ percentile for "low" density. However, the 90$^{\underline{th}}$ percentile for "high" density is 0.003 inch smaller than the 90$^{\underline{th}}$ percentile for "low" density. The effect is small and is primarily a slightly "broader" curve for the lower density condition. For this reason, although statistically significant, we attribute no practical difference in the density factor.

Table 4.4 Crack length percentiles (50 and 90) for significant factors
Values are calculated as averages across other factors. All values are in thousandths of an inch.

	level	angle	surface /facility	pos	flaw #	procedure/ facility	density
50th percentile	0	72	73	69	63	71	66
50th percentile	1	63	62	66	72	64	69
90th percentile	0	109	115	105	95	109	104
90th percentile	1	96	90	100	110	96	101

The accessibility of the lap splice, as measured by the low and high positions, also had a small effect. For the lower position, about 0.005 inch are added to the 90th percentile. Most of the inspectors commented that the lower row of lap splices in the experiment induced more discomfort than they would normally encounter. Besides the added discomfort, there was less light on the lower row due to the curvature. It is not unlikely that a combination of increased physical discomfort as well as increased task difficulties (locating and initiating the inspection at each rivet) affected the inspections the relatively small amount indicated.

The increased level of detection of the rivet sites with two cracks as compared to the rivet sites with a single crack is not inconsistent with a multiplicative model assuming independent chances of detecting each crack. The size of the detection level shift between inspecting rivet sites with single cracks to that of inspecting sites with two cracks is dependent upon the particular crack sizes.

4.4.2 Procedure and Surface Effects.

In specifying the factors to be modeled in section 4.4.1 it was noted that both the procedure and the surface effect were also reflective of facility effects. The procedure effects were not controlled in the collection of data. That is, each of the inspectors chose their technique for inspection. However, because of the overall balance in the choices between the sliding probe and the template technique, this effect was included in the PoD model of section 4.4.1.

The same number of inspectors used each technique, but the inspectors within a facility all tended to use the same technique. Thus, procedure and surface effects, as given in the PoD fit of section 4.4.1, could be reflective of facility differences unrelated to procedures or surface. In this section these effects in the presence of facility variations are explored in more detail. To facilitate the discussion, total variations are first broken down into within inspector variation (repeatability), between inspector variation, and facility variation components. The surface and procedure effects are then judged against the observed facility variation.

In section 4.2.2 curves for individual fits to each inspection were given. Individual points from those fits are used to estimate within inspector, between inspector, and facility variations. The estimated 50th and 90th percentiles from the individual PoD curves fit are used in the analysis. The full data table is given in Appendix B. (With the threshold model one inspector, F2, never achieved a 0.90 probability of detection. In this case, we substituted the 0.90 value obtained from the no threshold model.)

The data were also examined for highly influential points. Inspection F2 had substantial influence on inspection-to-inspection variation. Similarly, the inspections at facility G contributed greatly to estimated facility differences. Variance analyses with and without these factors are presented to gage their overall contribution.

At each facility one inspection was repeated. That is, the same inspector or inspection team did the complete inspection twice. In all cases the two inspections were separated by at least 3 days. This was done to minimize any learning effects concerning the nature of the test specimens. The test specimens were also shuffled between the inspections. Within each inspection, the basic layout regarding crack density factors was maintained, but the panels were shuffled within these groupings.

Each repeat inspection was done ostensibly under the same conditions. The same people, the same equipment, and the same general environment were repeated. Therefore the observed inspection differences are primarily reflective of day-to-day influences on the inspector or inspectors, including the inspector-equipment interface.

The repeat variation (within inspector), between inspector variation, and facility variation are estimated by an analysis of variance. The analysis of section 4.4.1 indicated that surface condition was a contributing factor. The surface condition is included in the model in order to not inflate the estimate of facility to facility variation. (The surface condition factor is examined in more detail later.) Facility variation is measured within each surface condition group. Inspector-to-inspector variation is measured within each facility and the repeatability is estimated from the 9 repeat inspections.

Table 4.5 shows the estimates as standard deviations in the *log*(crack length) scale where crack length is measured in thousandths of an inch. Also shown in parenthesis are the estimates that would result from eliminating Inspector F2 and all of facility G from the data. These values are given for comparison purposes, as it has been noted that these were influential inspections from the point of view of inspector variation and facility variation, respectively. More details are given in appendix B.

Table 4.5 *Estimated standard deviations of* $log(a_{50})$ *and* $log(a_{90})$
[·]-Inspection F2 and facility G removed.

Source	$\hat{\sigma}(log(a_{50}))$	$\hat{\sigma}(log(a_{90}))$
within inspector (repeatability)	0.077 [0.080]	0.095 [0.095]
inspector to inspector	0.127 [0.091]	0.215 [0.119]
facility to facility	0.090 [0.073]	0.105 [0.087]

Using the estimate of table 4.5 and probability values from the normal distribution, the estimate of "repeatability" can be characterized by saying that an inspector who nominally achieves a_{90}=0.100 inch (or $log(1000a_{90})$=4.605) could be expected to exceed a_{90}=0.113 inch (or $log(1000a_{90})$ =

4.605 + 1.282*0.095) approximately 10 percent of the time and similarly, would exceed approximately 0.117 inch approximately 5 percent of the time.

The variance estimates given above were calculated from analysis of variance tables. One can also test for differences in the PoD curves of the original and repeat inspections by appealing to the asymptotic normal properties of the maximum likelihood estimators. This method is explained in appendix D of reference 11. Four of nine repeat inspections (inspectors B4, F4, H2, and J1) are statistically different from the initial inspections (based on the no threshold model). There is no clear picture that the inspectors "learned" from the original inspection and therefore improved on the second inspection. There were also changes in the false call rates for many of the inspectors. A criterion shift internal to the inspector (as discussed in section 4.3) would explain these phenomena. Given the elapsed time (at least 3 days) criterion shifts could have occurred.

A variance components analysis, such as presented above, allows for a statistical test regarding the procedure and surface effects. For example, the data are grouped into two groups reflecting the surface condition of the skin specimens at the time of inspection. The difference in the mean levels is then compared to the estimated facility variation. The significance level reflects how likely the observed mean difference would be, if there was no real difference in the top level grouping. That is, one tests whether the observed mean differences could result from the inherent facility variation. This statistical procedure is applied to both surface effects and procedure effects separately. Significance levels are presented in table 4.6. Details are contained in appendix B.

Table 4.6 *Significance Levels for Surface and Procedure Effects [log(a_{50}), log(a_{90})]*

	Surface	Procedure
All data	[0.18, 0.19]	[0.25,0.70]
F2 and facility G removed	[0.31, 0.18]	[0.50, 0.87]

Table 4.6 indicates that the effects of surface condition and the procedures used are not statistically significant for the log(a_{50}) and log(a_{90}) values when compared to facility variation. We can also look at other values that reflect some aspect of detection curves, such as the background miss rate, C, that was estimated for each inspection. For this quantity there is no statistical evidence of facility differences. The facility-to-facility differences are no more than would be expected due to the between inspection variation. The average values of the background miss rate for each of the surface conditions and for the sliding and template initial procedures are shown in table 4.7. The estimated miss rate independent of crack size for Inspector F2 was 0.168. This value was more than twice any of the other inspections. With this atypical value removed, it is seen from table 4.7 that the two surface conditions differ by more than a factor of two (0.009 for the bare surface versus 0.022 for the painted). There is only a slight difference in the procedure averages, with the average on the template procedure being lower.

Table 4.7 Background Miss Rate (C) Averages by Surface and Procedures
[·]-Inspection F2 removed

	Surface		Procedure	
	Bare	Painted	Sliding Probe	Template
Average C	0.017 [0.009]	0.022	0.025 [0.018]	0.016

Thus, a surface effect that goes beyond that explained by facility differences is implied, whereas the procedure differences may be more facility related than due strictly to just the procedures.

Before leaving the question of the surface effect we look at an analysis of the data gathered from the full-size aircraft panels. One of these panels was painted and one was not. This condition was maintained at all nine facilities. Therefore a direct comparison can be made across all the facilities. Table 2.2 shows that the bare full-size panel had more sites with double cracks than did the painted panel. So as not to confound the effect of double flaws with the surface condition, curves were fit to the single crack data only. There was a significant difference in the two curves. Parameter fits are given in table 4.8.

Table 4.8 Full-Size Aircraft Panels Probit Model Fits for log(a)

Surface	μ estimate	σ estimate	C estimate	a_{50} (inch)	a_{90} (inch)
bare	4.1464	0.2994	0.006	0.063	0.093
painted	4.3316	0.2351	0.023	0.076	0.103

As does the data from the small panels, the data from the full-size aircraft panels indicates that there is a PoD curve shift to longer cracks for painted surfaces as compared to bare surfaces. The shift from bare to painted surface for the 50th percentile points of the PoD curve calculated from the full-size panels is 0.013 inch as compared to the average shift of 0.011 inch estimated from the small panels (table 4.3). Thus, the two sets of data are in substantial agreement for the 50th percentile. However, for the 90th percentile point there is a shift of only 0.010 inch from the full-size panels, as compared to 0.025 inch estimated from the data of the small panels.

The model used in section 4.4.1 estimated the factor effects without altering the estimated constant background miss rate, C. The analysis of the full-size panels indicates that this is not a good assumption for the surface condition. Overall, there is an approximately 0.010 inch shift in the PoD curves for inspections on painted surfaces as compared to those on bare surfaces, but there is also a shift in the background miss rate.

The change in the background miss rates from bare surface inspections to painted surface inspections could be partially due to increased task difficulty. That is, the task of locating the rivet head and aligning a probe to the rivet head is made more difficult when the rivet head is obscured by paint leading to more decision errors.

4.5 RETROSPECTIVE ANALYSIS.

In this section we will consider the techniques discussed in section 4.1.2. Using these techniques, we consider some of the factors for which data were gathered but which were not controlled within the experiment.

The data gathered on inspector backgrounds were considered as possible explanatory factors in an analysis of variance. The variables considered were age, eddy current experience, and recency of having done a lap splice inspection. These are the types of factors that often arise in explaining performance. Included with these factors were some of the factors that were considered in the probit analysis for the PoD function. These factors were surface condition, shift, team, and procedure (template, sliding probe).

The data on age and eddy current experience were in grouped categories. There were 3 age categories (20 to 30, 31 to 50, and over 50 years). Eddy current experience was coded 1 through 6 according to the inspectors' response on how often they performed eddy current inspections. The responses were grouped into daily (1), several times a week (2), weekly (3), several times a month (4), monthly (5), and less than monthly (6). The recency of the last lap splice inspection was given in weeks.

In the cases of the team inspections the average of the two age categories was used to describe the team. (When team members were in a different age group, they were in adjacent age groups.) For the eddy current experience and recency of lap splice inspection the minimum value between the two inspectors was used. For the inspectors who inspected twice, only the first inspections were considered.

A multivariate analysis of variance on the estimated 50th and 90th percentiles from the individual PoD curves was carried out. The analysis indicated that shift, team, and procedure were not significant in their contributions to explaining the variation of the estimated 50th and 90th percentiles. The analysis was rerun with these factors removed. The remaining factors (surface, age, eddy current experience, and recency of lap splice inspection) were all significant. (Details are in appendix B.)

However, almost all the inspectors were users of eddy current equipment either daily or several times a week. There was a single inspector for each of the categories 3 through 6. The one inspector, F2, who did rather poorly was a past supervisor who had just returned to regular inspection duty. This inspector had not been routinely performing eddy current inspections and was the sole inspector in category 6. Removing this one inspector and redoing the analysis eliminated the eddy current experience as an explanatory variable. However both age and recency of lap splice inspection were still significant.

The effect of the recency of lap splice inspection for explaining variation of the estimated 50th percentiles was marginal ($p=.04$), but the effect is dismissed for the following reasons. The estimated influence was in the direction of better performance with the longer times since performing the task. This runs counter to expectations. Figure 4.27 shows a graph of the estimated value of $log(a_{50})$ versus the number of months since the last lap splice inspection. From

figure 4.27 it is seen that the overall downward trend is influenced by one extreme point and a group of values at 24 months, which are mostly from the same facility.

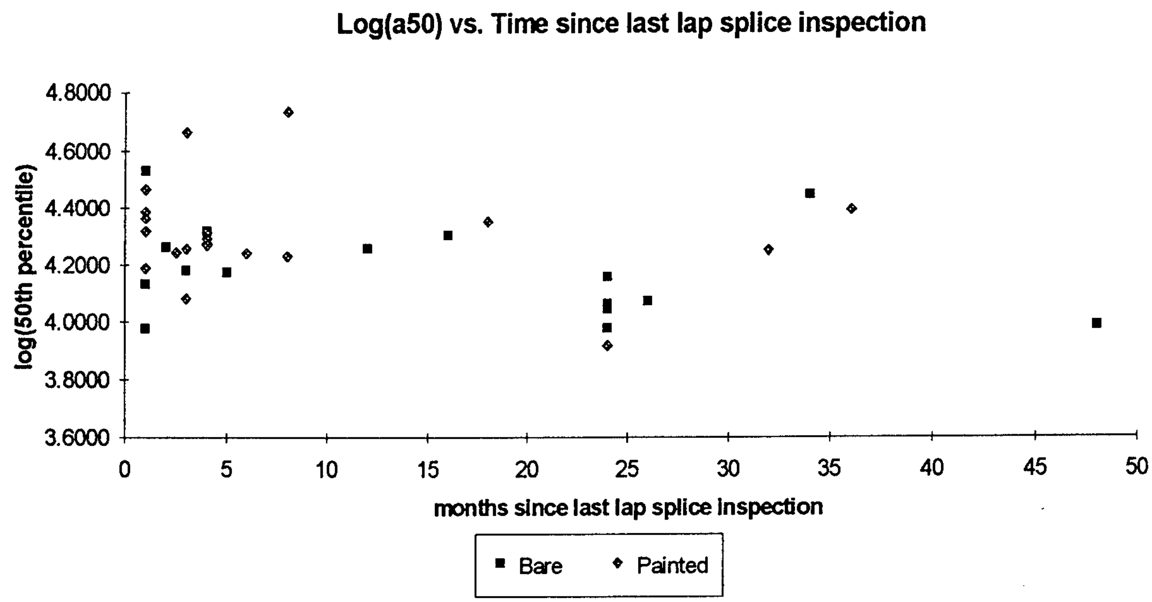

FIGURE 4.27 LOCATION FIT VERSUS LAP SPLICE INSPECTION EXPERIENCE (Better performance is at lower $\log(a_{50})$ values)

The age factor deserves some discussion. The estimates for $log(a_{50})$ and $log(a_{90})$ are both higher, on average, for the group of inspectors in age category 4 (over 50), than they are for the other categories. There are 8 inspectors in this category. Three of the inspectors are from Facility G, two are from Facility C, and one each from Facilities A, D, and J. Because Facility G was the poorest performing facility and most of its inspectors fall into the older category, we run the risk of attributing the variation to age when other facility specific factors are causing the differences.

It is worth noting the following relationships of the older inspectors' performances within their facilities: the inspector in Facility A was second best regarding performance, the two inspectors from Facility C were the worst and the best within that facility, the inspector in Facility D was the best, and the inspector within Facility J was the worst. (The younger inspector at Facility G performed the best within that facility.) Given that the facility variation is probably explained by a myriad of factors and the relationship of the older inspectors within the facility is not clearly in one direction, the indicated age effect should not be taken as conclusive.

In general, the various retrospective factors included in the analysis are not significant when the models incorporate facility effects. For example, the overall picture indicated that the age group affected performance. However, this overall pattern does not hold up within the various facilities. This may seem like an inconsistency, but is explained by the fact that many of the factors, such as age and experience, are more homogenous within facilities than between facilities.

5. DISCUSSION AND SUMMARY.

The Eddy Current Inspection Reliability Experiment grew from a basic question about how effective inspections were being done in the aircraft maintenance and inspection facilities. Various experiments had been performed in laboratory conditions, with inspectors that may or may not represent the field population of inspectors. Were inspections in the maintenance environment with all of its distractions capable of the same performance levels as had been demonstrated in a more benign laboratory setting?

The data gathered in this experiment indicate that the field factors, in aggregate, affect performance levels. From the laboratory inspections conducted as a baseline to the ECIRE, the 90th percentile point on the PoD for inspections on a bare surface is estimated to be in the 0.060 to 0.070 inch range. The same percentile (bare surface inspection) as an average from the field data is approximately 0.090 inch.

The ECIRE laboratory-derived PoD curves and an overall field-derived curve were presented in sections 4.2.1 and 4.4. They are shown again in figure 5.1 with the curves reported by Boeing in reference 4. The Boeing curves are shown because we believe that it is a natural question to ask how they compare. We do, however, caution against over-interpretation. Crack characteristics, such as orientation angles, or other known influential factors have not been accounted for in comparing the two experiments.

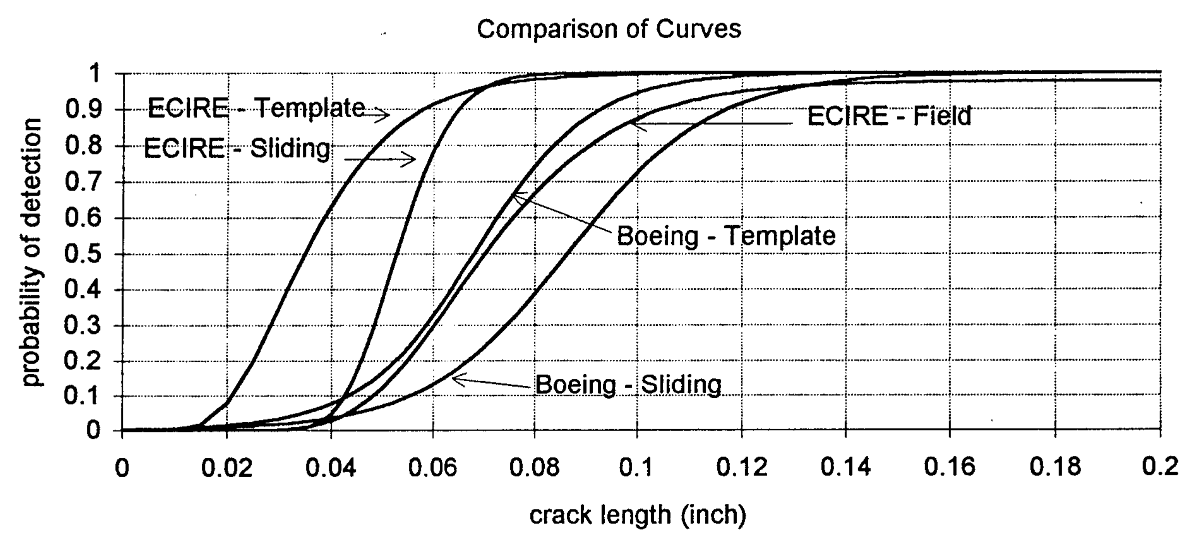

FIGURE 5.1 COMPARISON OF VARIOUS POD CURVES.
Boeing curves [4] were obtained using different test specimens and protocols.

5.1 INDIVIDUAL INSPECTION PERFORMANCE.

Procedures called for setting up the eddy current equipment on a calibration standard containing a 0.100-inch crack. Table 5.1 summarizes how each of the 45 inspections performed with respect

to this 0.100-inch level. Shown in the table for each inspection are the largest crack lengths missed and the number of sites missed that contained cracks exceeding 0.100 inch in length.

***Table 5.1** Individual Inspections -- Largest Crack Missed*

Inspection	Largest crack missed (inch)	number missed > 0.10 inch*	Inspection	Largest crack missed (inch)	number missed > 0.10 inch *
A1	0.158	1	F1	0.175	2
A1R	0.087	0	F2	0.812	19
A2	0.078	0	F3	0.118	1
A3	0.130	1	F4	0.222	1
A4	0.215	3	F4R	0.083	0
B1	0.075	0	G1	0.188	20
B2	0.322	5	G1R	0.240	17
B3	0.154	1	G2	0.096	0
B4	0.088	0	G3	0.322	25
B4R	0.174	3	G4	0.158	9
C1	0.087	0	H1	0.189	4
C2	0.192	1	H2	0.174	5
C3	0.160	2	H2R	0.812	2
C4	0.174	14	H3	0.322	7
C4R	0.174	5	H4	0.265	8
D1	0.102	1	J1	0.105	3
D2	0.169	2	J1R	0.098	0
D2R	0.118	2	J2	0.253	11
D3	0.083	0	J3	0.082	0
D4	0.085	0	J4	0.130	2
E1	0.102	1			
E2	0.186	5			
E2R	0.140	4			
E3	0.131	2			
E4	0.105	1			

* There were a total of 98 rivet sites containing cracks that exceeded 0.100 inch in length

Eleven of the 45 inspections (24 percent) were accomplished without missing any cracks over 0.100 inch in length. An additional 16 of the inspections (36 percent) missed one or two of the 0.100-inch cracks. On the other extreme, six inspections (13 percent) missed eleven or more of the 0.100-inch cracks. In the six extreme cases, other inspections at the same facility and using the same equipment were performed without any misses of cracks greater than 0.100 inch. Thus, factors specific to the inspectors and the procedures that they individually follow are implicated as a major source of variation, as opposed to inadequate equipment.

5.2 PROCEDURES.

In the laboratory environment distinct differences were measured for the sliding probe and template procedures. Unlike the laboratory results, the inspections in the field using the sliding probe procedure performed better than the template method. The choice of procedures was left to the inspectors, but there was a consistency of choice among the inspectors within a facility. Considering facility differences, the field results are best summarized by saying that those facilities choosing to implement the sliding probe procedure, on average, achieved better detection rates

than those facilities employing the template procedure. However, there is no information whether these facilities would have done as well or better had they used the template procedure.

Other facility factors could be influencing the sliding probe - template comparison. For example, three of the four facilities employing the sliding probe also used two person teams in performing the inspections. However, the results from the fourth facility indicate no significant difference in the probability of detection when using the team approach. The teams, on average, took about 20 percent less time to complete the inspections.

Some of the variation observed between facilities and between inspectors can be attributed to differences in procedures. Inspectors in three of the nine facilities departed greatly from the Boeing procedures. At two of the facilities, universal eddy current standard blocks were used to setup the inspection equipment. At the third facility no reference standard was used during setup.

The two facilities setting up on a depth notch from universal eddy current standard blocks each used a different width EDM slot. The performance in one facility was near the top while the performance of the other facility was among the worst. This illustrates that a departure from stated procedures does not in itself mean that a worse performance will result, but rather any departure from the developed procedures should be approached with care.

Boeing procedures call for a template verification of calls made by using the sliding probe or the rotating surface probe. This added step does not affect the probability of detection. It was overlooked by some of the inspectors, and when implemented, it was not uncommon for the template method to be applied using less sensitive instrumentation than had been used with the sliding probe.

Many inspectors did not establish a strict decision level for instrument signals based upon the calibration. Rather, they individually judged each signal from each inspection site. Some of these inspectors had better overall detection rates. Without an explicit criterion established during calibration, inspectors are more likely to mentally shift criterion levels during the inspection, thereby adding to the variability of results. A substantial criterion shift half way through the inspection task was observed in one inspector.

The false call rates for many of the inspections indicate that they were being performed at criterion levels set well above noise levels for uncracked sites. For these inspections it is not unlikely that relaxed criteria could be employed and would increase detection rates without an appreciable increase in false calls. This case was not, however, universal, and a relative operating characteristic analysis on some of the inspectors indicates that they were operating at a decision level where gains in detection rates would come only at the expense of more false calls.

5.3 TRAINING.

Although all the inspectors had received training in eddy current inspection techniques, many did not have equipment-specific training. The result of this lack of training was less than optimal equipment setups and lack of use of full equipment capabilities. Inspectors at one of the poorer

performing facilities were critical of management taking instrument-specific training and then doing an inadequate job of transferring that training.

A criticism that is often leveled against experiments of this nature is that the inspectors know that they are being tested and they know that the test specimens include cracks. This is true, but the inspectors still had to demonstrate a basic capability within their normal operating environment and based upon their training. There is no reason that the exhibited capabilities could not extend to actual aircraft inspections. All that is needed is the same attitude among the inspectors as they brought forth for this experiment.

5.4 POD MODELING.

The log logistic or the lognormal distribution functions provide similar fits to the data. Many of the individual inspections were better modeled by introducing a threshold parameter, C, to model a background miss rate. With this parameter in the model, PoD curves increase with larger crack lengths to (1-C) rather than to 1. The estimated background miss rates from the 45 inspections ranged from 0 to 0.168, with an overall average of 0.024.

The incorporation of the threshold, C, into the modeling is consistent with field observation of the inspectors. Some of the conditions contributing to background miss rates, independent of crack length, that were observed are inadequate procedures to maintain position when moving equipment, lapses of attention, intermittent problems with equipment, and the masking of audible alarms by ambient and intermittent noise in the facility.

Repeatability of inspections (the same inspector and equipment but different inspection times) is characterized by a standard deviation of $log(a_{90}) = 0.0948$. Consider an inspector who, on average, achieves a PoD of 0.90 at cracks 0.10 inch in length. Based upon the within-inspection variation estimated here, about five percent of the time we would expect the 90 percent detection rate to be achieved at cracks greater than 0.117 inch in length.

Inspector-to-inspector differences are a major source of variation in inspection results. Variation in the inspector population is characterized by a standard deviation of $log(a_{90}) = 0.2150$. Consider a population of inspectors who, on average, achieve a PoD of 0.90 at a 0.10-inch crack. Based upon this estimated variation, five percent of the inspectors would be expected to exceed 0.142 inch for a 90 percent detection rate.

Facility differences are significant. That is, the differences in facility performances, on average, were more than could be accounted for by inspector-to-inspector variation.

Other factors affecting inspections include surface condition, crack orientation, and "accessibility." Consider a base inspection as occurring on a bare surface, detecting horizontal cracks with good accessibility. Our estimate for the crack length at which there is 0.90 probability of detection (across inspectors) is 0.090 inch. For an inspection on a painted surface add 0.010 inch, but also expect an increase in a background miss rate independent of crack length. For off-angle cracks (in the 11 to 22 degree range) add 0.013 inch. For mild "accessibility" problems add 0.005 inch. In the context here, "accessibility" refers to an aggregate of the physical comfort level

of the inspector and the ease of performing the inspections (as could be influenced by such things as light levels).

Tedium, as measured by probability of detection differences over time, was not a significant factor. In general, the inspectors took adequate breaks and paced themselves in a manner to overcome the effects of tedium.

Individual variations in performance were not explained by the uncontrolled factors of age and recency of experience. These factors tended to have less variation within the facilities than existed between facilities. In the analysis these effects are therefore heavily influenced by facility differences.

5.5 CONCLUSIONS.

The following inspection process related conclusions result from the analysis of the field data combined with the observations of the monitors while the data were being gathered. There are no surprises within these conclusions. They are offered as a reminder to inspection facilities, NDI equipment manufacturers, and aircraft manufacturers as to the many influences on inspection reliability.

- *Individual inspector differences are a major factor affecting inspections.* No single individual factor related to experience and training was indicated in the observed data. However, it was common to find different inspectors within a facility exhibiting various levels of familiarity with the existing equipment. It is presumed that a better understanding of equipment capabilities would result in better inspections or at least less time required to perform the inspections. Training programs should be adopted that insure the individual inspector is knowledgeable about the specific equipment being used.

- *Environmental factors can influence an inspection.* Factors such as the accessibility of the inspection task and the condition of inspection surface affect the goodness of an inspection. The effects of any one environmental factor in isolation may not be large. By providing appropriate environmental conditions such as staging and lighting to ease the inspection task, better rates of detection will be realized. Inspection conditions such as lighting may assume a different importance depending on other factors. For example, the lighting may become more important when the inspection surface is painted due to an increased level of difficulty in locating the rivet heads.

- *Inspection misses result that are independent of crack length.* Inspectors and their supervisors need to be made aware of this phenomenon and to be especially cognizant of conditions that cause distractions and contribute to the miss rate. The data from this experiment indicate that background miss rates could be affected by procedural factors such as the condition of the inspection surface (painted versus bare). Some inspectors noted that they would rather the surface be stripped of paint before an inspection because they perceived an added difficulty in performing the inspection when the surface was painted.

6. REFERENCES.

1. Spencer, F., Borgonovi,G. , Roach, D., Schurman, D., and Smith, R., "Reliability Assessment at Airline Inspection Facilities, Vol. I: A Generic Protocol for Inspection Reliability Experiments," DOT/FAA/CT-92/12, I, March 1993.

2. Spencer, F., Borgonovi,G. , Roach, D., Schurman, D., and Smith, R., "Reliability Assessment at Airline Inspection Facilities, Vol. II: Protocol for an Eddy Current Inspection Reliability Experiment," DOT/FAA/CT-92/12, II, May 1993.

3. Norriss, T. H., "The Role of Non-Destructive Testing in Relation to Damage-Tolerant Aircraft Design," British Journal of Non-Destructive Testing, November 1985, vol. 27 no. 11.

4. Hutchinson, M. C., "Reliability of Eddy Current Lap Splice Inspection Procedures (Interim Report)," Boeing QCR&D document PE373A-1, April 1991.

5. Swift, T., "Information: Multi-Site Damage," FAA Memorandum from National Resource Specialist, Fracture Mechanics/Metallurgy to Manager, Aircraft Engineering Division AIR-100, January 3, 1991.

6. Samavedam, G. and Hoadley, D., "Fracture and Fatigue Strength Evaluation of Multiple Site Damaged Aircraft Fuselages - Curved Panel Testing and Analysis," Foster-Miller Report No. DTS-9024, May 1991.

7. Crosby, P. B., Quality is Free, McGraw Hill, New York, 1979.

8. Deming, W. E., Quality, Productivity, and Competitive Position, MIT Center for Advanced Engineering, Cambridge, Massachusetts, 1982.

9. "Eddy Current Inspection at Countersunk Fastener Locations," 737 NDT, Part 6, 53-30-05, Boeing Commercial Jet, Jan. 15/89.

10. Berens, Alan P., "NDE Reliability Data Analysis," *Metals Handbook Volume 17, 9th edition: Nondestructive Evaluation and Quality Control*, ASM International, 1988.

11. Annis, C., Berens, A., Bray, F., Erland, K., Hardy, G., Herron, W., Hoppe, W., "Proposed MIL-STD for USAF NDE System Reliability Assessment (Draft)," August 1989.

12. Hovey, P. W. and Berens, A P., "Statistical Evaluation of NDE Reliability in the Aerospace Industry," *Review of Progress in Quantitative Nondestructive Evaluation,* Vol. 7B, Thompson, D.O. and Chimenti, D.E., eds, pp. 1761-1768, Plenum Press, 1988.

13. Berens, A. P. and Hovey, P. W., "Statistical Methods for Estimating Crack Detection Probabilities," *Probabilistic Fracture Mechanics and Fatigue Methods: Application for Structural Design and Maintenance*, ASTM STP 798, J. M. Bloom and J. C. Ekvall, eds., American Society for Testing and Materials, 1983, pp. 79-94.

14. Swets, J.A. and Pickett, R.M., Evaluation of Diagnostic Systems: Methods from Signal Detection Theory, Academic Press, New York, 1982.

15. Swets, J.A., "Assessment of NDT Systems - Part I: The Relationship of True and False Detections." *Materials Evaluation*, 41:1294-1298, 1983.

16. Swets, J.A., "Assessment of NDT Systems - Part II: Indices of Performance." *Materials Evaluation*, 41:1300-1303, 1983.

17. Davis, M. K., "Proficiency Evaluation of NDE Personnel Utilizing the Ultrasonic Methodology," *eview of Progress in Quantitative Nondestructive Evaluation*, Vol. 7B, Thompson, D.O. and Chimenti, D.E. eds, pp. 1777-1789, Plenum Press, 1988.

18. Glasch, K., "Human Reliability in Nondestructive Evaluation," *Materials Evaluation*, 45:907-932, 1987.

19. SAS Institute Inc., *SAS/STAT® User's Guide, Version 6, Fourth Edition, Volume 2*, Cary, NC: SAS Institute Inc., 1989.

20. Hyatt, R.W., Kechter, G.E., and Menton, R.G., "Probability of Detection Estimation for Data Sets with Rogue Points," *Materials Evaluation*, November 1991, pp. 1402-1408.

Appendix A

EXPERIMENTAL PROTOCOLS

March 1993

Reliability Assessment Experiment

for

Eddy Current Inspection of Lap Splice Joints

in

Airline Maintenance and Inspection Facilities

Prepared by:

AEA Technology

England

CONTENTS

INSTRUCTIONS ON PROTOCOL USE

RAE 1

SCOPE

This document describes how the Protocols are intended to be used.

PROTOCOL USAGE

The complete set of Protocols are contained in this volume. They are presented in chronological order, as far as possible, so that they appear in the order in which they should be invoked in the experiment. They fall into two broad categories. The first includes Protocols which need not be replaced for each visit to a Facility and are classified as "permanent" or Reference Protocols. The second includes documents that have to be renewed for each visit. In particular, this latter category includes RAE 7 "Trial Checklist", and all forms. A form is defined as a document which the monitor is required to complete during the experiment, and a Table as one providing informative data to the monitor. A listing of the category of a Protocol and the number of forms required for each visit is given in Table 1.A, and this should used by the Custodian as a checklist when assembling the Protocol Folders for dispatch (see RAE 3).

RAE 7 is intended to be a working document which guides the monitors through each inspection, and which is also used to record information on equipment, environmental conditions, the Facility and each inspector. It is filed after each inspection and replaced with a fresh copy.

It should be noted that the schedule for testing inspectors at each Facility will likely be established at the time the monitors arrive. The monitors will be required to discuss the planned schedule with management during the briefing interview. If a monitor schedule has been established during pre-visit briefings, then the monitors will need to confirm or amend it as necessary.

TABLE 1.A

This list contains all the procedures and Forms required for the experiment. The Reference procedures are stored in the Permanent Folder, the Facility and Inspector documents are stored, initially, in the Facility Folder; after each inspection the forms for that inspection are filed separately in an inspection folder.

Custodian's Documents	Reference Procedures	Forms required for each Facility (No. of)	Forms required for each Inspector (No. of)		
	1			RAE 1	Instructions on Protocol use
	1			RAE 2	Inspector Supervision
1	1			RAE 3	Start Out
1		1		RAE/3/A	Packing Checklist
1				RAE/3/B	Equipment Shipping Record
	1			RAE 4	Management Briefing
		1		RAE/4/A	Inspector Schedule
	1			RAE 5	Start Up
	1			RAE 6	Reference Standards Experiment
		1		RAE/6/A	
		1		RAE 7	Trial Checklist
			1	RAE/7/Ins	Trial Checklist Form
			36	RAE/7/A	On-line comments
			8	RAE/7/B	On-line comments
	1			RAE 8	Panel Layout
			1	RAE/8/A	Panel Layout Checksheet
	1			RAE 9	Inspector Briefing
			1	RAE/9/A	Inspector Agreement
			1	RAE 10	Pre-Trial Questionnaire
	1			RAE 11	End of Trial Debriefing
			1	RAE/11/A	End of Trial Debriefing
	1			RAE 12	Data Recording and Transfer
	1			RAE 13	Close Down
		1		RAE/13/A	Facility Characteristics
	1			RAE 14	MOI Inspection
				RAE 15	Personal Log Book

INSPECTOR SUPERVISION PROCEDURE

RAE 2

SCOPE

One purpose of this document is to describe, as far as possible, the role and activities of the monitor in the field so that all parties understand and agree the actions required at an early stage. It is expected that the monitor will be familiar with the requirements on him at the time of the field trials. It also describes the course of action to be followed for variations in the planned program which might arise due to factors such as malfunction of equipment, mistakes, sickness, etc. Guidance is given on the type of comments required from the monitor on the inspector's performance, physical and mental skills, and personal characteristics relevant to inspection. It indicates how these subjective assessments can be quantified as far as possible. The document also contains a summary of the actions required of the monitor in data recording and document control.

CONTENTS

1. INTRODUCTION

In the reliability assessment experiment being conducted at selected aircraft inspection facilities in the USA by SAIC, qualified inspectors will each be required to examine test assemblies simulating lap splice joints in a Boeing 737 fuselage. Each inspector will participate in the experiment during his specified shift, using his normal inspection procedures, which should be based on Boeing Documents 737 D6-37239, part 6, subjects 53-3-03 and 53-3-05. His activities will be controlled, supervised and observed by the SAIC monitor managing the field tests. It is intended that the observation of the inspector's working characteristics should be as unobtrusive as possible, consistent with the objective of obtaining the required information on personal performance. Therefore, the monitor should position himself such that, as far as possible, the inspector is not disturbed by his presence. During the tests the monitor will be required to make specific observations and ensure that the experiment protocols and the declared inspection procedures are adhered to. This document gives guidance on the protocols to be followed.

It is planned to have two monitors on duty, one of whom will be an NDT expert and the other a human factors expert.

2. DESCRIPTION OF MONITOR'S DUTIES

The monitors should remember that they are not part of the facility organization and have neither the duty nor the right to enforce work rules, safety rules (except in cases of imminent death or serious injury) or administrative rules. They are there to observe specific behaviors in a specific setting.

2.1 MANAGEMENT OF TEST PROGRAM

The first duty of the monitor is to ensure that the test program proceeds as planned. There are a series of activities involved in this which are summarized below, with references to relevant detailed protocols, to give an overall activity chart for the visit to a facility.

i) The first act of the monitor upon arrival is to make contact with management immediately to arrange a meeting which will cover a detailed management briefing (RAE 4), agreement on the identities of the inspectors to be included in the tests and ensuring their availability at the required times, and arranging for the provision of a test area and appropriate supporting accommodation. He should also arrange to be supplied with the relevant calibration blocks at an appropriate time for the Reference Standards Experiment, and ask what inspection procedure will be used. If this is not the standard Boeing procedure then he should arrange to receive a copy of the procedure that the inspector will be working to, for archiving.

ii) The monitor is responsible for assembling the test frames and attaching the test panels in the correct sequence, and for corroborating that they are in the correct order. (RAE 8).

iii) There are several possible points at which the reference standards experiment can be performed, however it is recommended that the measurements are made before the first inspection. The reference standards experiment shall be performed in accordance with RAE 6.

iv) The monitor should summon the inspector for the tests at a time during the latter's period of work such that all the tests and interviews can be completed within the shift period. The first action will be to brief the inspector on the aims of the work and what he is required to do (RAE 9), and then complete the pre-trial questionnaire (RAE 10).

v) The activities of the monitor during the inspections are covered in the Trial Checklist (RAE 7), and include ensuring that the tests are performed according to the project protocols and the inspection procedures, making records of environmental conditions and the test equipment used, recording and checking the results (RAE 12), conducting the de-briefing interview (RAE 11), commenting on the performance of the inspector (see below), and rearranging the panels for the next tests as required by the panel layout procedure (RAE 8). He also is required to record any significant changes in the environmental conditions that occur during the tests.

vi) The activities defined in sections iv) and v) above are repeated for each inspector. When the entire test program is completed, the monitor has to disassemble the frame assemblies and load them, and the test panels, ready for exportation.

vii) The final act is to inform management that the work is over, answer any questions on the tests without giving away important information, and thank them for their cooperation.

2.2 INSPECTOR ASSESSMENT

The second duty of the monitor is to provide written comments giving his views on the capability and performance of the candidate (RAE 7). To ensure that the comments made on the inspectors cover those characteristics relevant to this assessment in as consistent a way as possible, given the subjective nature of the comments, this document gives guidelines for evaluating the main characteristics. A form has prepared for completion by the monitor, a copy of which is included in RAE/7/Ins. This does not preclude additional comments being made, should the monitor wish to do so, and any such comments should be written in the log-book at the time they arise.

2.3 ASSESSMENT OF PHYSICAL CONDITION

The third duty is to assess the inspector's physical condition during the tests. This covers his alertness, energy and behavior in the test conditions. These aspects are included in the Trial Checklist, RAE 7, and are explained in Section 4 below.

2.4 CONTROL OF DATA SHEETS AND DOCUMENTS

The fourth duty is to control the documentation and results sheets, and the coding and filing systems are described in Section 5. The monitor is responsible for recording, with the assistance of the inspector, the results declared by the inspector. It is planned that the data will be entered directly into a computer database, and the procedure for this and for confirming the entries is detailed in RAE 12.

2.5 SPECIMEN SECURITY

The fifth duty is to be in charge of the security and confidentiality of the specimens. The following precautions should be taken whenever the specimens have to be left unattended (by the monitors):

a) Small Panels: remove the panels from the frame and store in the locked shipping crate(s).

b) Large Panels: cover with tarpaulins and tie down with the bungee cords.

3. ABNORMALITIES OCCURRING IN PLANNED PROGRAM

The abnormalities covered are:

3.1 Slow inspections/late arrival
3.2 Equipment failure
3.3 Power failure
3.4 Sickness of inspector
3.5 Sickness of monitor
3.6 Non-arrival of inspector
3.7 Uncooperativeness
3.8 Mistakes
3.9 Program deviations
3.10 Damage to specimen panels

3.1 <u>Slow Inspections/Late Arrival</u>

The rate of working will have been derived from dry-run time checks, and the schedule allocated will be a reasonable average erring on the slow side. However, where it is clear that the test is going to over-run the end of shift then steps have to be taken as soon as the problem emerges.

	Extent of problem	Solution
3.1.1	**Moderate slowness.** For the case of a probable extension of about 1/2 hr on the shift	Attempt to carry on. If not possible see 3.1.3 below.
3.1.2	**Significant slowness.** Where the probable extension beyond the end of the shift is going to be about 1 hr	Discuss with the shift supervisor and the inspector the possibility of continuing. If this is not possible, see 3.1.3 below.
3.1.3	**Extreme slowness.** For a probable extension beyond the end of the shift	Arrange to complete the test on the next shift, or as soon as possible.

3.1.4	**Late arrival.**	This can be treated in the same way as slowness.

3.2 Equipment Failure

Eddy current equipment: Allow the facility inspector to change equipment for items as close as possible to the original, and record changes in daily log-book. Continue the test program with new equipment at an appropriate overlap point.

3.3 Power Failure

If this affects the tests, notify the local management and get expert help.

3.4 Sickness of Inspector

Notify local management and arrange for a re-run or a replacement at a suitable time in the future.

3.5 Sickness of Monitor

Get the back-up monitor to run the program, and if the sickness is likely to be prolonged contact base for a replacement.

3.6 Non-arrival of Inspector

Treat as for sickness, Section 3.4 above.

3.7 Uncooperativeness

This could have a serious effect on the experiment. The cause of dissent should be established and discussions taken with the local management and the inspector with a view to resolving the issue. In the limit however the candidate may have to be released from the program and a replacement obtained.

3.8 Mistakes

The definition of a mistake is an error made by the inspector which would result in him carrying out work which is not in the planned program and which could invalidate the test. For example, the use of equipment which is outside the specification, or the use of invalid procedures, or significant deviation from the procedures, etc. The latter case is dealt with in Section 3.9 below. The correctness of the gain setting used by the inspector is difficult to check at the outset, but the monitor should watch to ensure that the inspector is making some marks on the test piece while scanning the first row of rivets. If none or few are marked the monitor should ask the inspector to re-calibrate at the end of the row, as if it were part of the planned program. This provides the opportunity to investigate the cause of the low recording rate. When mistakes are noticed by the monitor the exact nature of the mistake and the consequences should be recorded in the log-book.

The definition does not include an incorrect result or errors such as non detection of flaws, incorrect amplitudes, or flaw positions, etc., since these results are at the center of the exercise to determine reliability.

3.9 Program Deviations

If the inspector is working in a manner which does not conform to the experiment protocols and which jeopardizes the achievement of the targets of the project, then the monitor must advise the inspector of the proper procedure.

3.10 Damage to Specimen Panels

Seven additional small panel specimens will be manufactured with the experimental group of 36 specimens, and will be maintained as backups.

If there is an extreme change in the expected pattern of inspector responses for a given specimen, then a change in crack characteristics may have occurred. If a change in the specimen is suspected, the panel should be sent back for characterization and a new one put in its place. The new one will not have exactly the characteristics of the old, so the changed pattern must be logged and noted in the data files.

In the case of loss or destruction of test specimens, the monitor should seek advice from the overall experiment manager.

4. ASSESSMENT OF INSPECTOR CHARACTERISTICS

Inspector assessment is covered by the Trial Checklist, RAE 7, which contains a copy of the data sheet to be filled in by the monitor at the end of each test. In some cases the factor being assessed is described in the Checklist by word opposites giving upper and lower limits. Between these limits are five boxes scaled from left to right, 5 down to 1. The monitor is asked to place a tick in the appropriate box. These markings will be used to normalize comments and also to give a semi-quantitative evaluation of each inspector.

For assessment of other factors, a description is required, for instance of job performance, scanning technique and posture. For job performance, the monitor is asked to observe if the procedure is followed. If this is not the case and the test is likely to be jeopardized then he should act in accordance with Section 3.9 above. If the deviation is minor then this should be noted in the log-book and transferred to the Checklist at the end of the test.

There is a general section in the Checklist entitled Other Influential Factors, which deals with the assessment of aspects such as disturbances, interruptions and unplanned breaks.

As a broad general assessment of confidence in the capability and reliability of the candidate the monitor is asked to decide whether he would employ the man on an inspection shift team.

5. CONTROL OF DATA AND DOCUMENTS

5.1 Coding System

5.1.1 **Inspector.** An alpha-numeric coding system will be applied by the monitor to the data records and documents. The method of allocating the identities is described in detail in Section 1.1 of Protocol RAE 7. In outline, they are formed by giving the Facility an alphabetic code (M, N, P, R, S, T, U, V, W), including the number of the test layout to be given to the inspector, allocating a number to each inspection at a given facility from 1 to 5, (which is dependent only on the order in which he is tested), and adding an identifier for the type of inspection i.e. E for an eddy current inspection and M for an MOI inspection. Where an inspector is performing a repeat inspection this will be indicated by a 'r' after the panel layout number.

An example is given below.

The first inspector at Facility M would be M101E(or M)

where M = facility (M,N,P,R,S,T,U,V or W)
1 = panel layout number
01 = # of the inspector at this facility (1-5)
E = EC technique
M = MOI technique
1, 2= Inspector designator when there are two people on the inspection team.

The panel layout which will be given for each inspection is identified in Table 7.A or 7.B of RAE 7.

5.1.2 **Specimen Panels**. The small panels are labeled A1, A2 etc., and the large panels are 101 and 102.

5.1.3 **Fasteners**. In any given test layout each fastener is uniquely identified by whether it is an **U**pper or **L**ower row, the **B**ay it is in, the **P**anel number in the bay and the **F**astener number in the panel:

UBiPjFk or LBiPjFk

where:

U = upper row
L = lower row
i = bay #, 1-6
j = panel #, 1-3
k = fastener #, 1-20

in all cases, counting is from left to right.

5.2 Control of Results Sheets, Protective Tape and Questionnaires

The filing system for documents and results sheets consists of a separate file for each inspector/inspection and one for facility related matters. Thus, there will be a set of 6 files for each facility, plus five more where an MOI inspection is to be carried out. These files will be given the alpha-numeric codes described above. Each length of protective tape will be marked at the end of each inspection, before it is removed, with the inspector's ID and the specific panel number. In addition the monitor will outline every other rivet and insert the rivet number (from left to right). The tapes will be stuck onto the white cards provided which are then inserted in Card Folder.

The monitor is also responsible for entering all relevant data into the computer data-base, according to RAE 12.

5.3 Filing

The monitor is responsible at the end of the inspection for filing all results and data sheets, and backing up the computer data.

START OUT

RAE 3

SCOPE

This procedure controls the packing and movement of the equipment crates. Two forms are used:

RAE/3/A provides a checklist of all the equipment required for the experiment. This is required by the custodian at the outset and by the monitors at each Facility.

RAE/3/B is used to control the transport of the crates to and from the storage hangar at Albuquerque. It is therefore required only by the custodian.

BACKGROUND

The first column in Form RAE/3/A has to be completed by the Custodian at Albuquerque when the crates are despatched to the first Facility and afterwards only if the contents are changed or the crates have been broken into. Subsequently, the Custodian need only complete Form RAE/3/B. The monitor is required to complete the last three columns in Form RAE/3/A at each Facility.

The Custodian is responsible for compiling the Protocols and Forms for each Facility. These are contained in two types of Folder. The Permanent Folder remains unchanged throughout the experiment, the Facility Folder is renewed by the Custodian for each visit. The copies of the Protocols which the monitor can use repeatably are included in the Permanent Folder (contents specified in Form RAE/3/A). The Forms and Protocols which will be completed at each Facility and returned to Mike Ashbaugh (as specified in RAE 13) are to be contained in the Facility Folder (as specified in Form RAE/3/A), and so this Folder needs to be packed by the Custodian in the transport crates for each visit.

PACKING CHECKLIST

FORM: RAE/3/A

FACILITY #

DATES

The five columns are:

1. No. of: The number of items that should be packed.

2. Initial Pack: Certifies that the correct number of items has been packed before setting out to the first Facility.

3. Unpack on Site: Refers to the checking of the equipment when it arrives at the Facility.

4. Pack on Site: Certifies that the correct number of items has been packed before leaving the Facility.

5. Load: Refers to the loading of the equipment onto the transportation to ensure that nothing is left at the Facility.

SPECIMEN HARDWARE	No. of	Initial pack	Unpack	Pack	Load
Triangular end supports	10				
5' wide test frame sections	6				
Small test panels	43				
Large test panels	2				
Pins	34				
Tubes	12				
Velcro straps	50				

☐ FACILITY FOLDER	No. of	Initial pack	Unpack	Pack	Load
This checklist Form RAE/3/A	1				
Inspector schedule RAE/4/A	1				
Reference Standards Experiment Form RAE/6/A	2				
Trial Checklist Forms RAE/7/Ins	6				
Inspection comments Form RAE/7/A	216				
Inspection comments Form RAE/7/B	48				
Inspector Agreement Form RAE/9/A	8				
Pre-trial questionnaire RAE 10	8				
De-briefing questionnaire RAE/11/A	8				
Facility characteristics Form RAE/13/A	1				
Empty folders for inspection results	6				

PERMANENT FOLDER	No. of	Initial pack	Unpack	Pack	Load
Guide to use of protocols: RAE 1	1				
Inspector Supervision: RAE 2	1				
Start Out	1				
Management briefing: RAE 4	1				
Start Up: RAE 5	1				
Reference Standards Experiment: RAE 6	1				
Trial Checklist Procedures RAE 7	1				
Panel layout RAE 8:	1				
Inspector briefing	1				
End of trial debriefing: RAE 11	1				
Close Down: RAE 13	14				
MOI Inspection: RAE 14	2				
Personal Log Book: RAE 15	1				

Label all the folders and forms with the Facility code and inspection number (if applicable)

COMPUTER EQUIPMENT (1)	No. of	Initial pack	Unpack	Pack	Load
Computer					
Power supply					
Printer					
Printer Paper					
Floppy disks	6				

Label the floppy disks with the inspection numbers (if known). In addition, use the DOS 'label' command to identify the floppy disks with the inspector number.

TEST INSTRUMENTS (1)	No. of	Initial pack	Unpack	Pack	Load
Eddy Current set					
Eddy Current probe					
Master reference standards	2				

MISCELLANEOUS INSTRUMENTS	No. of	Initial pack	Unpack	Pack	Load
Thermometer					
Humidity meter					
Light meter					
Noise meter					
Video recorder (at selected sites)					

(1) These items will travel with the monitors.

MISCELLANEOUS	No. of	Initial pack	Unpack	Pack	Load
Keys for the shipping crate locks					
Piano Dollies	2				
9/16" Open-Box end Wrenches	2				
9/16" Socket Wrenches	2				
5/16" Socket Allen Wrenches	2				
Rubber Mallets	2				
Socket Wrench Ratchets	4				
Small Pry Bar	1				
Wide putty knife (thick blade)	1				
3-foot carpenter's level	1				
Teflon Lubricant Spray (can)	1				
Paint					
Rolls of protective tape (3M No.336) (144yd x 2")	6				
Marker pens for marking indication on the tape	6				
Panel cleaning materials: alcohol, ammonia free cleaner, cloth					
Scissors					
Pens / pencils					
Floppy disk mailers					
A4 envelopes					
Stamps					
Shippers Airbills					
Fedex Mailers					
Spare Nuts and Bolts					
Chalk line					
Box knife					
Tarpaulins (16'x10')	2				
Bungee cords					
Protective tape results white card 20"x13"	30				
Folder for above item	1				
Packing Tape	2				

EQUIPMENT SHIPPING RECORD

FORMS RAE/3/B

This form is used:

i) to provide a record of the dispatch and receipt by the Custodian of the equipment crates, and

ii) to ensure that the locks on the crates have not been tampered with during transit or storage.

Facility	Locks OK Before	Date Sent	Date Received	Lock OK After	Comments

MANAGEMENT BRIEFING

RAE 4

SCOPE

This procedure is divided into two parts. Part A is intended to be read to relevant management personnel by the monitor upon arrival at an inspection Facility to brief them on the aims and objectives of the POD exercise. By stressing the importance of the work, it is anticipated that the full cooperation of the local management will be gained. The briefing session also allows management to clarify any questions they may have concerning the work.

Part B contains a list of the points that the monitor must resolve with management at this time.

The monitor should arrange a meeting with management as soon as possible upon arrival at the Facility.

PART A
NOTES FOR BRIEFING MANAGEMENT

Following the Aviation Safety Act of 1988, the FAA was directed by Congress to initiate an Aging Aircraft Program. The FAA has commissioned Sandia National Laboratories to design an experiment to determine the probability of detection (PoD) for cracks in aircraft components. The specific inspection chosen for evaluation in the experiment is the high frequency eddy current inspection of aircraft lap splice joints as covered by AD 88-22-11 R1 and AD 91-06-06. In addition, some inspectors will do inspections using Magneto-Optic/Eddy Current (MOI) equipment. SAIC has also been contracted to manage and conduct the field research in this experiment.

The test program is intended to evaluate both the technical capability of the eddy current inspection procedures and the equipment, as well as the human-factors issues associated with performing this inspection. The test program also is intended to evaluate the performance of inspections under industrial conditions, similar to those occurring on a routine basis. Professionally qualified inspectors familiar with working in the field are to be used throughout the tests.

For this experiment, a set of panels has been made which are representative of a Boeing fuselage. The panels contain lap splice joints in which some of the material around the rivets is defective. The inspectors are asked to inspect these panels using the techniques covered by the relevant Boeing procedures. The inspection requirements are those for skin lap joint inspections using high frequency eddy currents. The inspection requirements are covered by Airworthiness Directive 88-22-11 R1 (Boeing Service Bulletin 737-53A1039, Revision 5) and Airworthiness Directive 91-06-06 (Boeing Service Bulletin 727-53-0072, Revision 5). We are requesting that your inspectors perform the inspections on these test assemblies using normal equipment and procedures that would be used on a Boeing aircraft.

We plan to include four of your inspectors. To do this, we will need access to each inspector at an appropriate time during his work period for six to eight hours. The test setup contains a total of 77 feet of lap splice (924 inspection sites). We request that a work card be prepared for the inspectors participating in the experiment. The work card should specify the experimental setup as the inspection area, but otherwise should contain the appropriate procedural call outs and time estimates that would be present on work cards reflecting the same activity on an actual aircraft. During the inspection, the inspector will be expected to work to the experimental protocols (concerning the order of inspection and the reporting of results) so that the project objectives can be achieved. These will be explained to him by the monitor. The first inspector tested will be asked to repeat the tests at the end of the test cycle in some days time.

Any inspection that runs over the end of the shift will be continued on that inspector's (or team's) next regular shift.

Each inspector will be asked to sign that he/she agrees to participate in the experiment. This form will not be linked to the results in any way.

It is important for you to understand that the inspection results will be held absolutely confidential, and in no way will specific results be identified with your Facility or with the inspectors who take part. SAIC and Sandia are not obligated to provide the Federal Aviation Administration any direct linkage of inspection results to specific inspectors or specific facilities. As part of the implementation contract, SAIC has been tasked with assuring that inspector and facility confidentiality be maintained. To ensure this, we will wipe clean all references to you or your inspector at the end of the project, and we guarantee, since we control the data, that it will remain confidential. The confidentiality procedures are in place to protect both facilities as well as to safeguard human subject rights according to Federal regulations. The realization of the confidentiality process will necessarily mean that inspection results will not be available to the individual facilities.

Do you have any questions?

(The monitor should record any important questions in his personal log-book - RAE 15).

PART B
POINTS TO BE RESOLVED

1. After answering any questions, the identities of the inspectors, their availability and time of appearance should be confirmed with reference to Form RAE/4/A. If management wishes to change the schedule the monitor should amend the Form accordingly. He should agree with management that the inspectors and their supervisors will be informed of the inspection schedule.

The monitor must agree with management which shift will be duplicated and that the first inspector will do a repeat inspection at the end of the experiment.

Advice to Monitor

The inspection test schedule should have been agreed with the Facility management before the monitors arrive. If this has not been done or if changes are necessary the following advice is relevant.

Selection of Repeat Shift

The intent is for the participating inspectors to be reflective of the range of capabilities typically employed in inspection activities. Shift work is likely to reflect such attributes as experience levels and recentness of eddy current inspections; it is for this reason that shift work is considered to be a factor in the experimental design.

It is assumed that facilities employ inspectors on all 3 shifts, and therefore data will be gathered for all 3 shifts (3 inspectors). The shift for the 4th inspector is unspecified: at each facility the 4th inspector should be chosen from the shift employing the greater number of inspectors. If this is not possible he should be chosen to facilitate the logistics of the site visit, minimizing the burden on the monitors and the host facility.

If the facility only employs inspectors on two shifts, two inspectors should be chosen from each shift.

The layout/shift assignment is volatile and will be re-examined after 5 sites have been visited in order to determine if adjustments are needed in the remaining 4 sites to obtain a better balance.

Selection of Repeat Inspector:

In general, the first inspector at a facility (regardless of shift) should be the inspector who is requested to repeat his inspection at the end of the site visit. Thus, no attempt will be made to control which shift the repeat inspector represents. The intent is to have some passage of time so that the inspector is unlikely to draw on his memory of the first inspection results. The layouts for the two inspections will be slightly different and this will also minimize any memory effects. Logistics at each site may necessitate some modification of this plan however.

2. Determine whether a separate briefing to labor union management is required.

3. Establish what safety rules apply to the allocated area.

4. Where MOI is to be included, this should be discussed and the method of proceeding agreed.

5. Determine with management where the shipping crates are being held.

6. Inquire what Reference Standards are normally used for this type of inspection and arrange to have access to them for the Reference Standards Experiment (RAE 6) at the beginning of the work.

7. The facilities required by the monitor should be spelt out and agreed, e.g. workplace, power supply, desks, telephone, etc.(*), and the need to walk through the assigned location at an early stage should be pointed out. The purpose of the walk-through would be to verify the environment, relevant safety rules and any union rule requirements.

(*) These should have already been done in advance of arrival.

INSPECTOR SCHEDULE
FORMS RAE/4/A

FACILITY #

DATES

Shift		Mon.	Tues.	Wed.	Thurs.	Fri.
1	Name					
	EC or MOI?					
	Time					
2	Name					
	EC or MOI?					
	Time					
3	Name					
	EC or MOI?					
	Time					

START UP

RAE 5

SCOPE

This document gives guidance on the operations required upon arrival at a new Facility covering discussions with management and setting up the equipment. The stages beyond this are covered by separate protocols.

1. BRIEF MANAGEMENT

Perform the site management briefing as specified in RAE 4, which requires you to find out, in particular, where you can set up the experiment and details of the schedule for the inspectors.

Arrange to have access to the Facility's lap joint calibration specimens for the Reference Standard Experiment.

2. SITE PREPARATIONS

The equipment should have been delivered to site and be in place beforehand. If not, move the equipment to the area where the experiment will be held. Unpack the crates, checking each item off on the checklist Form RAE/3/A.

Inspect the panels and the support structures for any damage. Record any damage in the Log Book and refer to RAE 2 (Section 3.10) for guidance on the action to take.

3. FRAME ASSEMBLY

Note: For the assembly of the large and small frames a working space of approximately 12 x 12 yards is required, although some adjustment is possible, for instance the area could be 8 x 18 yards.

3.1 LARGE PANEL ASSEMBLY

A. Lay out the end support legs so that the two items for each leg are in a straight line, and so that the red spot color codes near one end of each item are adjacent. There are four legs for the large panel assemblies. For each leg assembly, fasten the two items together with a nut and bolt. Lay each leg so that it can bend upwards and secure the shorter leg section in the vertical position.

B. Set the separation of two of the end support legs to be the correct distance apart for one of the large panels. This is achieved using the wooden spacer bar provided with the assembly kit.

C. Take one of the large panels and place it face upwards across two of the legs so that one of the blue spots on the panel coincides with the blue spot on one of the legs. Fasten the panel to each of the end support legs.

D. Move the end sections of each leg upwards so that they join, and complete the bolting operation to make the support structure rigid.

E. Both monitors should combine to "roll" the complete assembly into the functional position; the large panel should be in an inclined vertical plane.

F. Repeat steps A. through E. for the second panel, and position it in line with first assembly. The direction of the line should be selected with the disposition of the finished small and large panel assemblies in mind.

G. **CHECK** that the edge of the lap joints in each of the panels faces downwards.

3.2. SMALL PANEL ASSEMBLY

A. Lay out the seven pairs of support legs as in i) above and bolt each leg section into a rigid triangular section, as described in iv) above.

B. Take the first end support leg and attach two of the tubular sections to the 4 in. stubs on the leg. Fix tubes in place with pins provided.

C. Attach another end support leg to the tubes by means of the 4 in. stubs and secure with pins.

D. Position the frame assembly along a line to be back-to-back with the large panel assemblies, and consistent with the space available.

E. Attach two tubes to the 4 in. stubs on the assembled frame support section, and repeat step C.

F. Repeat step E. until the six bay sections are assembled.

G. Using the adjustable feet on the supports and a spirit level, ensure that the lower section of each end support is horizontal in two directions.

H. Using a device such as a water level or strong elastic, ENSURE that the tubular frame members are in line and at the same height; the latter is achieved by adjusting the height of the stub support plates.

I. Attach the small panel support frames to the tubular members and fix in position with the Velcro straps from the kit provided. From the front of the assembly, ENSURE that the frames line up as well as is practically possible. Time should be devoted to this activity.

The small panel frame assembly is now complete. The small panels will be attached prior to the inspection as specified in RAE 8.

4. EQUIPMENT CHECKS

Check any monitoring equipment to ensure that it is working, e.g.:

Noise meter	Computer
Light meter	Printer
Thermometer	EC set
Humidity meter	

5. REFERENCE STANDARDS EXPERIMENT

Perform the Reference Standards Experiment:

GO TO PROCEDURE RAE 6 - REFERENCE STANDARDS EXPERIMENT

REFERENCE STANDARDS EXPERIMENT

RAE 6

SCOPE

This procedure details the steps involved in the Reference Standards Experiment. This experiment is designed to compare a Facility's lap-joint calibration block(s) with the master reference block. The procedure is based on the calibration stage of the oversize template method of inspection as specified in the Boeing Procedures 737 D6-37239, part 6 subjects 51-00-00 and 53-30-05.

EQUIPMENT

Reference EC Set: Elotest B2
Reference Probe: Make Model
Unpainted Master Reference Calibration Standard (Type 369)
Painted Master Reference Calibration Standard (Type 369)
Template with the hole to be used in the experiment clearly identified
Results Form RAE/6/A
0.002 - 0.004" (0.005 - 0.010 cm) thick non-conductive shim
Type 369 calibration standard(s) obtained from the Facility

PROCEDURE

Select the appropriate painted or unpainted master reference calibration standard for the experiment, depending on which type of calibration standard has been provided by the Facility.

STAGE 1

The first stage in the Reference Standards Experiment is to calibrate the reference eddy current inspection equipment using the master reference standard.

A. Put the probe on the surface of the master reference standard at least 0.5" away from the edge of the block and the cracks. Balance the instrument according to the manufacturer's instructions.

B. Adjust lift off to obtain less than 5% of full screen deflection when the probe is slid from the non-conductive shim to the bare surface of the standard.

C. Put the probe guide on the reference standard.

D. Visually center the marked hole in the template around the 5/32" rivet head with the crack.

E. With the probe guide held firmly in place, scan around the circumference of the rivet head. Monitor the instrument response. You should be able to clearly identify between the

sudden instrument response from the reference standard crack and the slow instrument response from an off-center condition.

F. Adjust the sensitivity control to obtain a 50 to 80% of full scale deflection when passing the probe across the crack.

G. Record on Form RAE/6/A the values for the three controls; Yx2, dB and ϕ. Also record the crack signal peak amplitude (% fsd). Record the other (unchanging) front panel settings in the comments section of the form.

NOTE: For the master reference standard block the settings should be:

Yx2 $\cong$ 3   dB $\cong$ 30 $\phi \cong 10°$ amplitude $\cong$ 50 - 80% fsd

If you are suspicious about the values you have obtained, record the fact in the Log Book.

STAGE 2

The next stage is to check the Facility's calibration standard(s). This is done by keeping the gain/sensitivity setting on the eddy current equipment fixed whilst the Facility's calibration standards are inspected.

H. Leave the inspection equipment set up as determined above from the master reference block.

I. On each calibration block within the facility visually center the marked hole in the template around the 5/32" rivet head with the crack.

J. With the probe guide held firmly in place, scan around the circumference of the rivet head. Monitor the instrument response. You should be able to clearly identify between the sudden instrument response from the crack and the slow instrument response from an off-center condition.

K. Record the crack signal peak amplitude (% fsd) on Form RAE/6/A.

STAGE 3

The third stage is to independently calibrate the Facility's blocks and record the settings for comparison with the inspectors' values obtained later. Avoid, if possible, changing any of the controls other than the three to be recorded. (See Stage 1, Item G)

L. Put the probe on the surface of the Facility's block at least 0.5" away from the edge of the block and the cracks. Balance the instrument according to the manufacturer's instructions.

M. Adjust lift off to obtain less than 5% of full screen deflection when the probe is slid from the non-conductive shim to the bare surface of the block.

N. Put the probe guide on the calibration block.

O. Visually center the marked hole in the template around the 5/32" rivet head with the crack.

P. With the probe guide held firmly in place, scan around the circumference of the rivet head. Monitor the instrument response. You should be able to clearly identify between the sudden instrument response from the crack and the slow instrument response from an off-center condition.

Q. Adjust the sensitivity control to obtain a 50 to 80% of full scale deflection when passing the probe across the crack.

R. Record on Form RAE/6/A the values for the three controls; Yx2, dB and ϕ. Also record the crack signal peak amplitude (% fsd). Record the other (unchanging) front panel settings in the comments section of the form.

REFERENCE STANDARDS EXPERIMENT

FORMS RAE/6/A

FACILITY #

DATE

TIME

BLOCK REF.	STAGE	Yx2,dB,ϕ	% fsd	COMMENTS
Master Reference Block	1			
	2	same as Stage 1		
	3			
	2	same as Stage 1		
	3			
	2	same as Stage 1		
	3			

Sheet ______ of ______

TRIAL CHECKLIST

RAE 7

This document is meant to be used by the monitors in preparing for and recording information specific to each inspection. It covers both the standard eddy current and MOI inspections. It starts with the preparations for each inspection and takes the monitor through to the end of an inspection session. Various other procedures are referred to by this procedure, they are :

RAE 2	Inspector Supervision Procedure
RAE 8	Panel Layout Procedure
RAE 9	Inspector Briefing Procedure
RAE 10	Pre-Trial Questionnaire
RAE 11	End of Trial Debriefing
RAE 12	Data Recording and Transfer Procedure
RAE 13	Close Down Procedure
RAE 14	MOI Inspection

This procedure is written as a guide to a check list to be filled in for each inspection. The instructions (RAE 7) and the check list elements (RAE/7/Ins) are separated in order to facilitate the data gathering process. The latter forms are to be filed following the site visits.

I - BEFORE INSPECTOR(S) ARRIVES

1. INSPECTION CODE

The monitor's first task is to allocate a unique code number to each inspection and inspector (in team situations). This code is defined to include all the relevant test parameters:

e.g. M101E(or M)1(or 2)

where M = facility (M,N,P,R,S,T,U,V or W)
1 = panel layout number
01 = # of the inspector at this facility (1-5)
E = EC technique
M = MOI technique
1 or 2 = Inspector designator in two person teams

Create the Inspector's code as follows:

1.1 Enter the Facility code letter
(M,N,P,R,S,T,U,V or W)

1.2 Select and enter the Layout number from Table 7.A according to the inspector's shift. If this is a repeat inspection on one of the inspectors then enter the layout number as #r.

1.3 Allocate a number to the inspection based on the chronological order at this facility; use a two digit code, starting at 01, 02, etc.

1.4 Enter "E" for a standard eddy current inspection.
Enter "M" if the inspection will be done using the MOI.

1.5 Enter a 1 or 2 to differentiate inspectors if a two person team is performing the inspection. This transform the inspection code (created up to this point) into an inspector code for use in differentiating the inspectors.

Code					
Instruction	1.1	1.2	1.3	1.4	1.5

1.6 Enter the Inspector's code on the front of

RAE 10
RAE/11/A

In addition, enter the inspection code at the top of each page of RAE/7/Ins (check list for this document) and Forms RAE/7/A and RAE/7/B (on-line comment sheets).

2. EQUIPMENT NEEDED

2.1 Check you have all the necessary hardware and materials for the next test:

- keys for shipping crate locks.
- cleaning materials (removing marks from panels)
- tools for rearranging panels
- protective tape
- markers for marking indications on protective tape
- Permanent Folder of protocols
- Data file containing pack of forms for next inspector:
  (Ensure all sheets are correctly labeled with the inspector's code and the date)

  Panel layout check sheet, RAE/8/A, or computerized check sheet
  Inspection check list RAE/7/Ins
  On-line comments during inspection: Forms RAE/7/A (36 sheets) and RAE/7/B (8 sheets)
  Agreement to participate, RAE/9/A
  Pre trial questionnaire RAE 10 Debriefing questionnaire RAE/11/A

- Computer
- Printer
- Printer paper
- Floppy disk (labeled with the inspection code and the date)
- Thermometer
- Humidity meter
- Light meter
- Noise meter
- Video recorder for MOI inspections

3. PANEL SET UP

The selection of the panel layout to be given to an inspector depends on:

- the facility
- the shift (day, evening or night)
- whether or not the inspector has already done a test inspection
- the type of inspection: conventional eddy-current or MOI

The order in which the panel assemblies are inspected depends on the panel layout.

There are 8 panel layouts:

1, 2, 3, 4, 1r, 2r, 3r and 4r

For each inspector the layout number will have been determined above from Table 7.A for the Eddy Current inspections or Table 7.B for the MOI inspections.

3.1 Attach the small panels to the frame assembly using the procedure RAE 8 (PANEL LAYOUT), and rearrange the positions of the large panels as shown in RAE 8.

3.2 Check that the inspection area is ready for the next inspector -i.e. marking pens available and that any previous marks have been removed from the panels.

II - WHEN THE INSPECTOR ARRIVES

4. INSPECTOR PREPARATION

4.1 Give the inspector a copy of the INSPECTOR BRIEFING Procedure (RAE 9) and read it to him.

Ensure the inspector understands what is required.
Ensure the inspector has signed the 'Agreement to participate'.

4.2 Ask the inspector to fill in the Pre-Trial Questionnaire (RAE 10).

4.3 Ask to see the inspector's procedures.

4.4 Record the reference number and date of the procedures and the level of approval on form RAE/7/Ins. Also record the basis of the procedure. The procedure may reference a Service Bulletin (SB) or Airworthiness Directive (AD). If so, record the reference given. If the procedures are internal to the facility, arrange to receive a complete copy and file this in the inspection file.

4.5. On RAE/7/Ins record the method specified in the eddy current inspection procedures for detecting and evaluating a defect signal: (This will be compared with what is actually done, as recorded later.)

4.6 Discuss the test protocol with the inspector to ensure that he understands the items which are specific to these tests:

- How to mark an indication.
- The assignment of a confidence level to each indication.
- The assignment of a confidence level if the indication has had to be confirmed using a different technique.
- The fact that surface scratches are not to be recorded.

4.7 If the inspector is unhappy about any aspects of the procedure, note on RAE/7/Ins. If the inspector says he cannot perform the inspection as required, refer to the Inspector Supervision Procedures (RAE 2).

5. CALIBRATION & RECORDING OF CHARACTERISTICS

5.1 Check that the calibration block brought by the inspector is one of the blocks you had access to for the Reference Standards Experiment. If it isn't, make a note of the number and perform the Reference Standards Experiment when the inspector has finished his tests.

5.2 Cover the calibration block with the protective tape.

5.3 Ask the inspector to start his calibration stage.

5.4 Whilst the inspector is calibrating, one of the monitors should complete the environmental characteristic section of RAE/7/Ins.

5.5 When the Inspector has finished the calibration, record the details about the equipment and the inspection technique to be used in RAE/7/Ins.

5.6 Record the calibration settings with the inspector on form RAE/7/Ins.

6. PERFORM INSPECTION

6.1 Ensure you have 36 Forms RAE/7/A (small panels) available, and 4 Forms RAE/7/B (large panels).

6.2 Select the order in which the panels should be inspected from the table below. Tell the inspector in which order to inspect the panels, and ask him/her to start the inspection.

Layout No.	Start	→	→	Finish
1, 1r	Upper	Lower	Unpainted	Painted
2, 2r	Painted	Unpainted	Upper	Lower
3, 3r	Lower	Upper	Painted	Unpainted
4, 4r	Unpainted	Painted	Lower	Upper

Note: Upper and Lower refer to the splice rows on the small panel assembly. Painted and Unpainted refer to the large panel assemblies.

6.3 Follow the progress of the inspections on Form RAE/7/A for the small panels and Form RAE/7/B for the large panels, recording the times at various fixed points e.g. the triangular supports, or for each panel. Record any other comments of a broader nature in the space headed "Other Influential Factors" on form RAE/7/Ins.

Include any comments you may wish to make. These might cover:

- the occurrence and duration of any breaks.
- any fasteners on which the inspector spent more time than usual.
- any instances when the inspector called for assistance from a colleague.
- any instances where you communicated with the inspector.
- any instances where the inspector went back to re inspect a fastener (other than those called for by the procedure)

Identify fasteners using the convention given in RAE 2 Section 5.1.
i.e.

UBiPjFk or LBiPjFk

where:

U = upper row
L = lower row i = bay #, 1-6
j = panel #, 1-3
k = fastener #, 1-20

in all cases, counting is from left to right.

6.4 During the inspection, record information about the inspector's general characteristics on RAE/7/Ins.

7. POST EXPERIMENT INTERVIEW

7.1 Interview the inspector using procedure RAE 11 and Form RAE/11/A (End of Trial Debriefing).

III - AFTER THE INSPECTOR HAS LEFT

8. RECORD INSPECTION RESULTS

8.1 Record the results of the inspection using the procedure given in RAE 12 (Data Recording and Transfer Procedures)

9. ADDITIONAL INFORMATION

9.1 Record your overall impression of the inspector using the appropriate section of RAE/7/Ins.

10. TRANSFER OF HUMAN FACTORS DATA TO THE COMPUTER

10.1 Transfer the data from the trial checklist and the facility's characteristics checklist into the computer. If any values were not recorded, measure them now. Print out the checklist.

10.2 Compare the manuscript and computer versions of the checklist. Correct any mistakes in the computer version.

10.3 Place the checklist in the inspector's file.

10.4 Backup the data obtained in the session.

11. DATA QUALIFICATION

11.1 Collate all the forms for this inspection:

Trial checklist (RAE 7 - this document)
On-line comments during inspection:
    Form RAE/7/A 36 sheets
    Form RAE/7/B 8 sheets Panel layout check sheet RAE/8/A
Agreement to participate RAE/9/A
Pre trial questionnaire RAE 10
Debriefing questionnaire RAE/11/A
Computer print outs of checklists

11.2 Check that all the forms are legible.

11.3 Ensure that the inspectors code number, date and time (where applicable) are on all the forms.

11.4 Place the forms inside the inspectors file. Ensure that the file is marked with the inspector's code number.

12. SPECIMEN PANEL DE-MOUNTING

12.1 Check that there are no marks on any of the panels; clean the panels if necessary.

12.2 Inspect the plates for any signs of damage or distinguishing marks. Note any such areas in the Log Book. If there are an areas which could affect the experiment, either because they are in the inspected row, or they are so obvious that they could be used to identify the panel, then refer to RAE 2.

12.3 Remove the small panels from the frame and place the panels in the shipping crates. If you are leaving the area temporarily, close and lock the crates and place the protective covers over the panel assemblies in accordance with RAE 2.

When all five inspections have been completed:

PROCEED TO PROCEDURE RAE 13 - CLOSE DOWN

TABLE 7.A v1.0

LAYOUT NUMBER FOR STANDARD EDDY CURRENT INSPECTIONS

NOTE: The layout/shift assignment is volatile and will be re-examined after 5 sites have been visited in order to determine if adjustments are needed in the remaining 4 sites to obtain a better balance. Check that you have the latest version of this table.

i) 3 Shift Facility

It is assumed that facilities employ inspectors on all 3 shifts. The shift for the 4th inspector is unspecified (U): at each facility the 4th inspector should be chosen from the shift employing the greater number of inspectors. If this is not possible he should be chosen to facilitate the logistics of the site visit, minimizing the burden on the monitors and the host facility.

Three Shift Facility Layout Numbers

Shift	Facility Code								
	M	N	P	R	S	T	U	V	W
1 (day)	1	4	3	2	1	4	3	2	1
2 (evening)	2	1	4	3	2	1	4	3	2
3 (night)	3	2	1	4	3	2	1	4	3
U	4	3	2	1	4	3	2	1	4

ii) Two Shift Facility

It the facility only employs inspectors on two shifts, two inspectors should be chosen from each shift.

2- Shift Facility Layout Numbers

Shift	Facility Code								
	M	N	P	R	S	T	U	V	W
1ST Day Shift	1	4	3	2	1	4	3	2	1
2ND Day Shift	3	2	1	4	3	2	1	4	3
1ST Night Shift	2	1	4	3	2	1	4	3	2
2ND Night Shift	4	3	2	1	4	3	2	1	4

Layouts for the MOI Inspections should parallel those for the eddy current facility.

ON-LINE COMMENTS - SMALL PANELS
FORMS RAE/7/A

Inspector #																				
Row	Upper / Lower																			
Bay	1 / 2 / 3 / 4 / 5 / 6																			
Panel	1 / 2 / 3																			
Fastener	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Com ment	a																			

Com ment	Time	Notes
a	:	Start of scan
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	

ON-LINE COMMENTS - LARGE PANELS
FORMS RAE/7/B

Inspector #	
Large Panel #	101 / 102

Fastener	Time	Notes
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	
	:	

INSPECTION CHECKLIST
FORM : RAE/7/Ins

This checklist is organized into two parts. The first part covers the gathering of information that is specific to the calibration and equipment being used. The second part is specific to the inspection environment and the inspectors. The numbers on the checklist correspond to the instructions given in RAE 7.

CALIBRATION AND EQUIPMENT

4.4 Record the reference number, date of the procedures, and level of approval

____

Is the procedure based on (check if applies):

a) 737 D6-37239, part 6, subject 53-30-03 _______
or b) 737 D6-37239, part 6, subject 53-30-05 _______
or c) other specification ______________________
or d) is the source not specified _________________

If the procedures fall into c) or d) , arrange to receive a complete copy and file this in the inspector's file. Check that the procedures are compatible with the test inspection:
e.g. Do they call for the surface condition to be different from that of the test panels?

If there are any discrepancies, note these below and discuss them with the inspector.

4.5 Briefly record below the method specified in the eddy current inspection procedures for detecting and evaluating a defect signal: (This will be compared with what is actually done, as recorded later.)

i) Recording threshold:

ii) Confirmation method and threshold specified in procedure:

5.5 When the Inspector has finished the calibration, record the details about the equipment and the inspection technique to be used .

Technique to be used (Check)	Template Method	
	Oversize Template Method	
	Sliding Probe Method (Confirmed with oversize template)	
	Other:	
EC Set	Manufacturer:	
	Model:	
	Serial #	
	Certification in date?	
Probe:	Manufacturer:	
	Reference #:	
	Contact / Non contact	
	Differential / Absolute / Other:	
	Pancake / Toroid Coil / Other:	
	Coil diameter (units)?	
	Shielding:	
Cables:		
Calibration Standard		

5.6 Record the calibration settings with the inspector .

Gain(s)	Horizontal: Meter: Vertical
Frequency (kHz)	
Filtering	
Surface	Painted / Unpainted (tape)
Calibration Level	units =
Inspection threshold	units =
Coil output impedance	
Digitization	
Template hole size	
Rotating probe speed	

TOOLS	
Availability (describe)	
Clean ?	
Employee furnished? (describe)	
Adequate? (describe)	
Calibrated? (describe)	

INSPECTION CHARACTERISTICS

4.7 Record any uneasiness of the inspector(s) about the procedures he/she is being asked to follow for the inspections.

5.4 Environmental Characteristics

TEMPERATURE (Check relevant box)					
Cold & Drafty	Cold	Comfortable	Hot	Very Hot	Hot & Humid
Variable?	Measured Value_______°F				

HUMIDITY (check relevant box)		
Dry	Comfortable	Humid
Measured Value _______%		

LIGHTING (Check relevant box on each line)				
Artificial	Mixed	Outdoors		
Glaring	Bright	Average	Dim	Shadowed
Measured Value_____________ft.candles				

NOISE (Check relevant box on each line)		
Steady level	Intermittent	
High	Acceptable	Low
Measured Value _________dBA		

GENERAL ATMOSPHERE (Check relevant box)					
Gloomy	High pressured	Noisy (see above)	Cheerful	Quiet	Relaxed

DURING THE INSPECTION

6.4 During the inspection, record information about the inspector's general characteristics:

PHYSICAL CONDITION						
Tired	1	2	3	4	5	Energetic
Unwell	1	2	3	4	5	Well
Distressed by condition	1	2	3	4	5	Comfortable in conditions

ATTENTIVENESS						
Gaze continually wanders while scanning	1	2	3	4	5	Always watches screen while scanning

POSTURE (standing, seated, kneeling, etc.)
a) Upper Row
b) Lower Row
c) Large Panels

SCANNING TECHNIQUE:
a) Scan speed (units)
b) Scan direction (e.g. L to R)
c) Other
Location of EC or MOI set with respect to inspector:

ENVIRONMENTAL CONDITIONS:

Are there any significant changes to the environmental conditions noted earlier during the calibration stage?

If so, amend the values.

JOB PERFORMANCE:

Describe any deviations from the inspector's own procedure. If these will affect the validity of the results, refer to RAE 2 and record the action taken. Cross reference any actions with the relevant Form RAE/7/A Sheet #.

Describe any occurrences which could affect the validity of the results. Record the action taken (refer to RAE 2). Cross reference any actions with the relevant Form RAE/7/A Sheet #

OTHER INFLUENTIAL FACTORS
Highlight any other significant activities that are related to the human factors elements: **SMALL PANELS:** Upper Row Lower Row **PANEL 101** **PANEL 102** MISCELLANEOUS

9. ADDITIONAL INFORMATION

9.1 Record your overall impression of the inspector using the questions below.

ATTITUDE						
Interested	5	4	3	2	1	Bored
Co-operative	5	4	3	2	1	Uncooperative
Relaxed	5	4	3	2	1	Tense
Hard working	5	4	3	2	1	Lazy
Well motivated	5	4	3	2	1	Disenchanted
Careful	5	4	3	2	1	Slipshod
Easy-going nature	5	4	3	2	1	Belligerent
Would you employ him/her on your team? Certainly / Probably / At a pinch / No						

WORK PATTERN (i.e. how the inspector organizes the tasks						
Worked conscientiously (did not waste time)	5	4	3	2	1	Wasted Time (performed unnecessary tasks)
Kept work area tidy	5	4	3	2	1	Work area always messy

PANEL LAYOUT

RAE 8

SCOPE

This procedure covers the setting up of the panel assemblies for each inspection.

1. EQUIPMENT

Tools for removing and attaching panels.
Panel container with panels.
Protective tape.
Sprung panel clamps.

2. SELECTION OF PANEL LAYOUT

Each inspection has been allocated a panel layout of small and large panels which is dependent on shift and Facility. The layout number will have already been allocated in RAE 7 for the incoming inspection. Use this layout identity number to determine the sequence of small and large panels for the next inspection from Table 8.A (small panels) and Table 8.B (large panels) in this Protocol.

3. PANEL ATTACHMENT

A. If applicable, unclip the small panels from the frame and store in the correct position in the panel storage container.

B. Working from left to right along the specified layout, select the panels one by one and clip them into position using the sprung clips provided, with the edge of the lap-joint facing downwards. It is important to avoid discontinuities at joints as far as possible.

C. Arrange the large panel assemblies as required by the layout specification.

D. When the panel arrangement is completed, move to the rear of the panel assemblies with the laptop computer. Using the layout checksheet, fill in each panel number versus its physical position. From the back of the small panel assembly start at the right-hand end of the top row and work to the left. Repeat this procedure for bottom row. The large panels have a number on the left-hand edge (from the back). From the rear the first panel is the one on the right. The computer check sheet will not allow wrong panel numbers to be entered. Therefore, any disallowed entries will require the relocation of the incorrectly arranged panels.

E. When assembling the panels for the first time at a Facility it is necessary for the monitor to clean the surfaces in a typical way. Use the cleaning fluid provided.

F. Attach protective tape to cover the top row of fasteners: the lower edge of the tape should be approximately halfway between the first and second rows of fasteners.

RETURN TO PROCEDURE RAE 7 - TRIAL CHECKLIST

TABLE 8.A Small Panel Layout Arrangements - All viewed from the front

LAYOUT 1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A39	A40	A1	A23	A2	A32	A3	A4	A24	A11	A12	A35	A13	A28	A14	A38	A15	A16
Lower	A41	A42	A6	A7	A31	A5	A26	A9	A33	A10	A18	A27	A17	A36	A22	A19	A29	A21

LAYOUT 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A41	A42	A6	A7	A31	A5	A26	A9	A33	A10	A18	A27	A17	A36	A22	A19	A29	A21
Lower	A39	A40	A1	A23	A2	A32	A3	A4	A24	A11	A12	A35	A13	A28	A14	A38	A15	A16

LAYOUT 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A41	A42	A16	A15	A38	A14	A28	A13	A35	A12	A11	A24	A4	A3	A32	A2	A23	A1
Lower	A40	A39	A21	A29	A22	A19	A36	A18	A27	A17	A10	A33	A9	A26	A5	A31	A7	A6

LAYOUT 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A40	A39	A21	A29	A22	A19	A36	A18	A27	A17	A10	A33	A9	A26	A5	A31	A7	A6
Lower	A41	A42	A16	A15	A38	A14	A28	A13	A35	A12	A11	A24	A4	A3	A32	A2	A23	A1

Table 8.A continued.

LAYOUT 1r

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A40	A39	A11	A24	A4	A3	A32	A2	A23	A1	A16	A15	A38	A14	A28	A12	A35	A13
Lower	A42	A41	A10	A33	A9	A26	A7	A31	A5	A6	A21	A29	A19	A36	A22	A17	A27	A18

LAYOUT 2r

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A42	A41	A10	A33	A9	A26	A7	A31	A5	A6	A21	A29	A19	A36	A22	A17	A27	A18
Lower	A40	A39	A11	A24	A4	A3	A32	A2	A23	A1	A16	A15	A38	A14	A28	A12	A35	A13

LAYOUT 3r

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A42	A41	A13	A28	A12	A35	A14	A38	A16	A15	A1	A23	A2	A32	A3	A4	A24	A11
Lower	A39	A40	A17	A27	A18	A21	A36	A22	A29	A19	A6	A7	A31	A5	A26	A9	A33	A10

LAYOUT 4r

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Upper	A39	A40	A17	A27	A18	A21	A36	A22	A29	A19	A6	A7	A31	A5	A26	A9	A33	A10
Lower	A42	A41	A13	A28	A12	A35	A14	A38	A16	A15	A1	A23	A2	A32	A3	A4	A24	A11

Table 8.B Large Panel Arrangements - All viewed from the front

Layout #	Left hand panel	Right hand panel
1, 1r	Unpainted	Painted
2, 2r	Painted	Unpainted
3, 3r	Painted	Unpainted
4, 4r	Unpainted	Painted

INSPECTOR BRIEFING

RAE 9

SCOPE

This procedure is intended for the use of the monitor immediately prior to a test session to inform the incoming inspector of the aims and objectives of the POD exercise, and to give a general introduction on what the tests involve and how they will be performed. The importance of the work will be stressed at this time in order to gain the inspector's cooperation. The briefing session also allows the inspector to clarify any questions he may have concerning the work. The monitor should make a note in the daily log of any points raised in discussion.

A copy of this procedure is to be gven to the inspector at the time of the briefing.

NOTES FOR BRIEFING INSPECTOR

Following the Aviation Safety Act of 1988, the FAA was directed by Congress to initiate an Aging Aircraft Program. The FAA has commissioned Sandia National Laboratories to design an experiment to determine the Probability of Detection (POD) of cracks in aircraft components. The specific inspection chosen for evaluation in the experiment is the high frequency eddy current inspection of aircraft lap splice joints as covered by AD 88-22-11 R1 and AD 91-06-06. In addition, some inspectors will do inspections using Magneto-Optic/Eddy Current (MOI) equipment. SAIC has been contracted to manage and conduct the field research in this experiment.

The test program is intended to evaluate both the technical capability of the eddy current inspection procedures and the equipment, as well as human-factors issues associated with performing this inspection. The test program also is intended to evaluate the performance of inspections under industrial conditions, similar to those occurring on a routine basis.

For this experiment, a set of panels has been made which are representative of a Boeing fuselage. The panels contain lap splice joints in which some of the material around the fasteners is defective. You will be asked to inspect these panels using your normal equipment and your usual procedures for lap splice inspections, working at your normal rate. The period of work will be during your normal shift. If you cannot complete the inspection during a single shift, you will be asked to finish it during your next regular shift.

You are asked to report any indication that exceeds the threshold reporting criteria specified in the procedures that you use for lap splice inspections. You should do this by drawing a ring (circle) around the fastener that shows the indication.

[The monitor should be sure to READ the next section to the inspector(s)]

In addition, as a separate exercise to help our data collection, we are asking you to say how "confident" you are in your call that the indication is reportable. That is, we want you to mark the circle with a 1, 2, or 3, indicating:

3 - means you are absolutely certain that the indication is reportable.

2 - means that you are reasonably sure that the call of an indication is correct.

1 - means that you have some doubts about the indication being reportable, but that you cannot overlook it.

It is important to note that you do not have to mark your confidence in the call when you circle the rivet. The confidence ratings are important to our analysis, but we want you to follow your own inspection procedures as far as possible. Thus, you can leave the decision about your level of confidence until later and return to the rivet to mark your confidence rating. You should, of course, mark any indication breaking the reporting threshold when you get that indication.

For example, at the beginning of an inspection - you may have doubts about what confidence rating to give your call. You can skip putting down a confidence rating at that time and come back to it later. If you go back to a marked fastener to give it a confidence rating and you think that it is no longer a valid call, then you should mark it with a zero. I can show you this in this

drawing. (Show Figure 1) In this example, one fastener has been marked confidently with a 3, and the call at another fastener has been changed to a no-call after further examination by marking out the former rating and putting in a zero. If you think that the indication is caused by a surface scratch, then do not mark it at all.

One last thing - if you call an indication, we want you to mark the position (where the crack is) and the orientation of any cracks at the fastener, as indicated by the signal.

> **IF YOU ARE USING A ONE-STEP TECHNIQUE, LIKE THE OVERSIZE TEMPLATE TECHNIQUE, THE NEXT INDENTED PARAGRAPH DOES NOT APPLY TO YOU.**

> **IF YOU ARE USING A TWO STEP TECHNIQUE, SUCH AS A SLIDING PROBE, FOLLOWED BY A VERIFICATION TECHNIQUE, READ THE NEXT INDENTED PARAGRAPH.**

> The method of marking fasteners and confidence ratings has been explained to you. However, because you are using procedures requiring verification, you will have to do the marking in two stages. In the first scan along the row, you are asked to mark a ring around any fastener which is found to have a reportable indication, as I showed you in the drawing. You then return to it later, to examine it with the verification technique. In this case, you should mark the confidence rating at the time of verification - although you may defer giving the rating until you have had a chance to look at a few verifications, if you so desire.

Do you have any questions at this stage on how to mark indications?

You will be required to do some eddy current inspections over a painted surface and some over an unpainted surface. For each of these cases you should use your normal inhouse procedures and calibration methods.

You have been chosen for this work because you are professionally qualified and familiar with working with the Boeing procedures in routine eddy current inspections of lap splices. However -

it is important that you understand that this is not a test of your own personal skills, and that all results and analysis data will be held absolutely confidential regarding your name or how you personally performed. Your name will never be linked with any specific data and data will only be reported in statistical summary form. To ensure this, we will wipe clean all links between these data and your name and the name of your facility at the end of the project. We guarantee, since we control the data, that it will remain confidential.

The inspections will be managed and observed by monitors whose role is to ensure that the inspections are performed according to the overall experimental plan, and to record relevant data and results on the environment and inspection conditions. Our aim is to gather statistics on the probability of crack detection, and in no sense will we be making judgements on you or any of the inspectors taking part in the experiment.

We ask you not to discuss your results with any of the other inspectors, since this could affect the studies and make it difficult to obtain unbiased data from the other inspectors.

You will be asked to sign a form to say that you agree to participate in the experiment.

Do you have any questions?

Notes for Monitor:

1. For the first inspector at each Facility inform him that he will be required again for a second session at the time and date already agreed with his management.

2. Some inspectors will be asked to apply the MOI technique. Where appropriate, tell the inspector that he will be expected to use this technique and the conditions applying.

3. Check the procedures to be used and if they are not the standard Boeing procedures obtain a copy and file them.

4. When the inspector understands what he/she has to do, and is happy with the arrangements, ask him/her to sign the Agreement to Participate, Form RAE/9/A.

O O O O O O O O O O
O O O O O O O O O O
O O O O O O O O O O

Figure 1. Details of how to mark indications.

INSPECTOR AGREEMENT

FORMS RAE/9/A

AGREEMENT TO PARTICIPATE

I have read and heard the explanation of research purposes given above. I grant permission for Sandia National Laboratory and SAIC and their designated agents to observe me at work with the test pieces and to obtain information about training, experience and other relevant information. I understand that my participation in this research is strictly voluntary. I understand that I may withdraw my participation in this research at any time without any effects on my job standing or employee status. I understand that the information about me is strictly confidential and will not be used in anything but statistical summary form.

SIGNED: .... DATE: ....

PRINT NAME: ....

PRE-TRIAL QUESTIONNAIRE

RAE 10

INSPECTOR #	SEX	DATE	TIME

SCOPE

The aim of this questionnaire is to acquire information on human factors aspects of the inspector.

The questionnaire should be completed by the inspector immediately prior to his test session.

A. GENERAL

1. What is your Age:

Under 20	20 - 30	31 - 50	Over 50

B. EDUCATION

2. How would you classify your general educational level:

(Please check as appropriate)

High School or highest grade completed	
Attended College	
College Degree	
Technical Degree	

C. CAREER PROGRESS

3. Which of the following are true for you?

(you may check more than one)

I expect to be still working in aircraft inspection in five years time	
I expect to be at a higher grade within two years	
I am interested in a career change	
I am properly qualified	
I plan to achieve further qualifications	

4. Please summarise your career history with approximate dates of significant changes

Date	Company	Job Title	Responsibilities

D. TRAINING

5. How much training have you had on the equipment you will be using today? (Please give details)

6. When was your last training or refresher on this equipment? (Please give details including approximate dates).

7. Was your training on this equipment or similar equipment provided in the classroom or as on the job training from an experienced user.?

E. EXPERIENCE

If you are performing an MOI inspection please answer the questions with respect to the MOI.

8. How often do you use Eddy Current (MOI) equipment?

9. How often do you inspect lap-joints with Eddy Current (MOI) Techniques?

10. How long has it been since your last lap-joint inspection?
(Days/weeks/months, approximately)

11. What do you think the priority of NDI is compared to the other things that you need to do?

- To you personally:

- To your management:

F. GENERAL CONDITION

12. How well did you sleep last night/day?

Badly ←→ Very Well

13. What time did your shift start today?

14. What task(s) have you been doing prior to coming to these tests?

15. Compared with normal, please rate how you feel at the moment with respect to the following:

	less		same		more
1 Irritability					
2 Efficiency					
3 Depression					
4 Sociability					
5 Physically tired					
6 Mentally tired					
7 Alertness					
8 Feeling healthy					
9 Thinking ability					
10 Happiness					

16. How do you feel about being asked to take part in this experiment?

END OF TRIAL DEBRIEFING

RAE 11

SCOPE

At the end of each eddy current inspection it is important to record the inspector's perception of how well his/her work went. This will be achieved by the monitor leading the inspector through a structured interview. The questions posed in the interview are listed in the Form RAE/11/A, attached, and this form should be used by the monitor to record the answers.

PROCEDURE

The structured debriefing interview should take place after the inspector has cleaned himself up, and be conducted in a relaxed way in the work place.

Thank the inspector for his/her participation and ask him/her if they would mind answering a few questions about themselves and their impressions of the experiment to help to improve future research of this type. Encourage the inspector to feel that he/she has stepped outside the experiment and is now helping the monitor to assess the effectiveness of the experiment. Make sure that as you question the inspector you do not make judgemental statements or show your opinion of his performance.

For example; it is acceptable to ask

"what happened when ...."

but not

"what went wrong ..." or "what was the problem ..." etc

The way in which you ask a question, your tone of voice or your facial expression can show your perception of his performance, whether good or bad, and this will be noticed by the inspector and may reduce his level of co-operation. Let him be the one to state that he had problems or that everything went really well.

END OF TRIAL DEBRIEFING

FORMS RAE/11/A

INSPECTOR #

DATE

TIME

A. INSPECTOR WELL-BEING

1 Do you feel fit and well at the moment?

YES	
NO	

2 Do you feel tired?

YES	
NO	

B. TECHNICAL

3. How much was working on this test arrangement like working on an aircraft? (Please check one box)

Very	
Somewhat	
Similar in some ways, not in others	
Somewhat dissimilar	
Very dissimilar	
Not at all similar	

4. What were the major differences between these inspections and actual aircraft inspections. Please describe the three most major below.

a)

b)

c)

5. Do you think that the differences described in Question 4 (above) affected the inspection results? If yes, please explain.

6. How would you suggest that we improve this work?

7. How much did the watcher bother you?

A lot	
Some	
Not at all	

8. Were there any parts of the work that you had difficulties with. If so what were they?

9. Were there any parts of the work which you found particularly interesting, unusual or challenging. If so what were they?

10. Now you have finished the experiment, what is <u>your</u> interpretation of the confidence markings you gave?

C. GENERAL CONDITION

11. Ask the inspector which of the following describes how he felt overall during the tests.

Ask for reasons for unusual answers.

	Interested	Bored	
	Relaxed	Tense	
	Irritated	Calm	
	Fresh	Tired	
	Sleepy	Awake	
	Attentive	Inattentive	
	Enjoyed the work	Disliked the work	

Reasons:

DATA RECORDING AND TRANSFER

RAE 12

SCOPE

This procedure covers the recording and transfer of the inspection results from the marks on the specimens.

PROCEDURE

The markings on the panels should be stored in two separate ways:

i) Forms RAE/7/A and RAE/7/B: The inspector's markings should be transferred to the relevant form. The completed forms should be stored with the other forms for this inspector in the inspector's file.

ii) Protective tapes: These should be labelled with the panel position (i.e. bay number (1-6) and panel number (1-3)) and the inspector's code number. One of the monitors should then mark and number every other fastener before removing the tapes and sticking them onto the white card. The white cards should be stored in the folder provided.

Subsequently, the results should be entered into the computer database.

CLOSE DOWN

RAE 13

SCOPE

This document covers the operations involved in preparing to leave a facility at the end of all inspections.

1. FACILITY CHARACTERISTICS QUESTIONS

Record on Form RAE/13/A your general impressions of the way the facility operates.

2. DEBRIEF MANAGEMENT

Perform the site management debriefing as specified in RAE 2 (Section 2.1.vii). Concentrate on thanking the management. Avoid discussing the results; your position is that you do not know where the cracks are so you cannot say how well any of the inspectors have done. If pressed you can say that they worked very professionally.

3. DISASSEMBLE FRAME ASSEMBLIES

Dismantle the frame assemblies and inspect the structure for any signs of damage. Record any damage in the log book; refer to RAE 2 (Section 3.10) for guidance on what action to take.

4. EQUIPMENT CHECK

Pack the frames and end supports in the shipping crates.

Check that all the facility specific forms are labelled, completed and present in the facility file. These are:

- equipment checklist (RAE/3/A)
- inspector schedule (RAE/4/A)
- reference standards experiment results (RAE/6/A)
- facility checklist (RAE/13/A)

Refer to the checklist (RAE/3/A). Check you have all the necessary hardware and materials.

Lock the crates.

5. SHIPPING

Arrange for the equipment to be shipped to either the next site or the storage site.

When the equipment is loaded for shipping, use the checklist RAE/3/A to ensure that all the equipment leaves site.

Send the hard copy files and floppy disks to M. Ashbaugh using the Fedex box. Do not post using the facility's mail system.

Leave the site.

FACILITY CHARACTERISTICS
FORMS RAE/13/A

FACILITY #

DATE

TIME

MANAGEMENT	
High pressure? (describe)	
Schedule constraints? (describe)	
Describe intercommunications	
Team cohesion? (describe)	
Distracting tasks? (describe)	
Does the Facility have a busy workload?	

HOUSEKEEPING	
Neat?	
Well organized?	
Describe unusual characteristics (good or bad)	

Appendix B

Data Fits and Statistical Analysis Details

LIST OF FIGURES

LIST OF TABLES

In this appendix various parameters fits that are referred in the text are given. Various other background data are also given. All table entries of estimated parameters are consistent with crack lengths being expressed as thousandths of an inch (mils). That is, a crack of length 0.100 inch would be entered as 100 mils. Graphs are labeled accordingly.

Parameter fits for individual inspections

Table B1 contains the maximum likelihood estimates for $\mu = -\alpha/\beta$ and $\sigma = 1/\beta$ and the threshold (background miss rate), c. Also given are the estimates of μ and σ when c is not included in the model. The last two columns are for the no threshold model fits for μ and σ. If the threshold maximum likelihood fit was 0, the fits are the same as those given in the first two columns and are not repeated.

Table B.1 Lognormal Maximum Likelihood Estimates by Inspection

	lognormal parameters		Threshold	No Threshold lognormal	
Inspection	$\hat{\mu}$	$\hat{\sigma}$	C	$\hat{\mu}$	$\hat{\sigma}$
A1	4.2571	0.1247	0.010	4.2402	0.2275
A1R	4.2922	0.1337	0.000		
A2	4.1742	0.1000	0.000		
A3	4.1798	0.1199	0.010	4.1804	0.2079
A4	4.1580	0.0565	0.021	4.1420	0.3428
B1	4.0828	0.2016	0.000		
B2	3.9016	0.2454	0.048	3.8026	0.6863
B3	4.2282	0.1846	0.010	4.2215	0.2599
B4	4.2757	0.1885	0.000		
B4R	4.1138	0.1926	0.027	4.0967	0.3981
C1	4.2644	0.1342	0.000		
C2	4.1311	0.1849	0.009	4.1060	0.3102
C3	3.9733	0.2353	0.015	3.9639	0.3631
C4	4.5340	0.3249	0.000		
C4R	4.4084	0.1401	0.043	4.4326	0.2790
D1	4.3057	0.2512	0.000		
D2	4.3149	0.2031	0.013	4.3120	0.2735
D2R	4.2274	0.2682	0.000		
D3	4.0736	0.3161	0.000		
D4	4.0623	0.3129	0.000		
E1	4.3521	0.1272	0.000		
E2	4.3078	0.2443	0.034	4.3266	0.3750
E2R	4.3582	0.2824	0.000		
E3	4.2897	0.1299	0.018	4.2928	0.2099
E4	4.2423	0.3339	0.000		
F1	3.9702	0.3116	0.024	3.9582	0.4653
F2	4.3849	0.2423	0.168	4.5266	0.6806
F3	4.0439	0.1150	0.014	4.0564	0.2440
F4	3.9847	0.2723	0.009	3.9541	0.4043
F4R	4.0701	0.1916	0.000		

table continues

Table B1. (continued)

	lognormal parameters		Threshold	No Threshold lognormal	
Inspection	$\hat{\mu}$	$\hat{\sigma}$	C	$\hat{\mu}$	$\hat{\sigma}$
G1	4.6624	0.2962	0.010	4.6708	0.3082
G1R	4.5864	0.1788	0.073	4.6529	0.3522
G2	4.3940	0.1493	0.000		
G3	4.7109	0.2808	0.066	4.7861	0.4068
G4	4.2534	0.5672	0.000		
H1	4.2348	0.1757	0.038	4.2267	0.3848
H2	4.3059	0.2065	0.032	4.3201	0.3121
H2R	4.2141	0.2358	0.012	4.1089	0.5348
H3	4.3731	0.1627	0.068	4.3819	0.4881
H4	4.2557	0.4319	0.024	4.2688	0.5125
J1	4.4666	0.1561	0.000		
J1R	4.2717	0.2227	0.000		
J2	4.3306	0.2853	0.086	4.3974	0.4817
J3	4.2566	0.1271	0.000		
J4	4.1877	0.3426	0.000		
ALL	4.2434	0.290474	0.023	4.2466	0.3923

Comparison of Model Fits with and without Threshold

For selected inspections, Figures B.1 through B.3 graphically compare the threshold model to the probit model without the threshold. The largest effect in including a threshold parameter is to keep the front part of the curve (PoD < .5) in a region that more adequately reflects the data. The threshold fits give a more optimistic view of the crack length at which a .9 probability is achieved, but at the expense of perhaps never reaching a higher probability level.

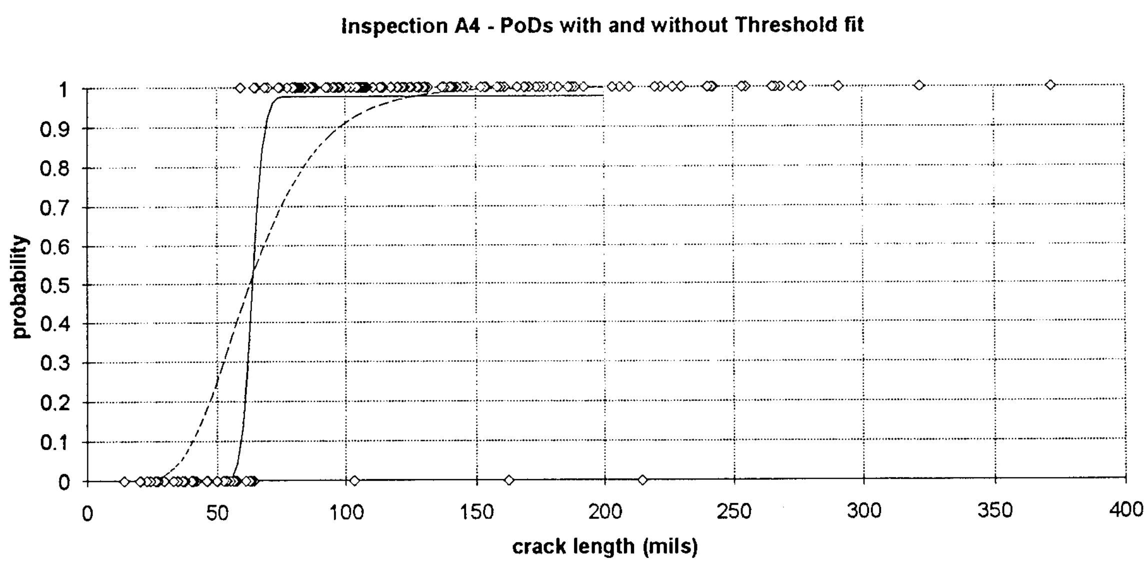

FIGURE B.1 THRESHOLD AND NON-THRESHOLD MODELS FOR INSPECTION A4

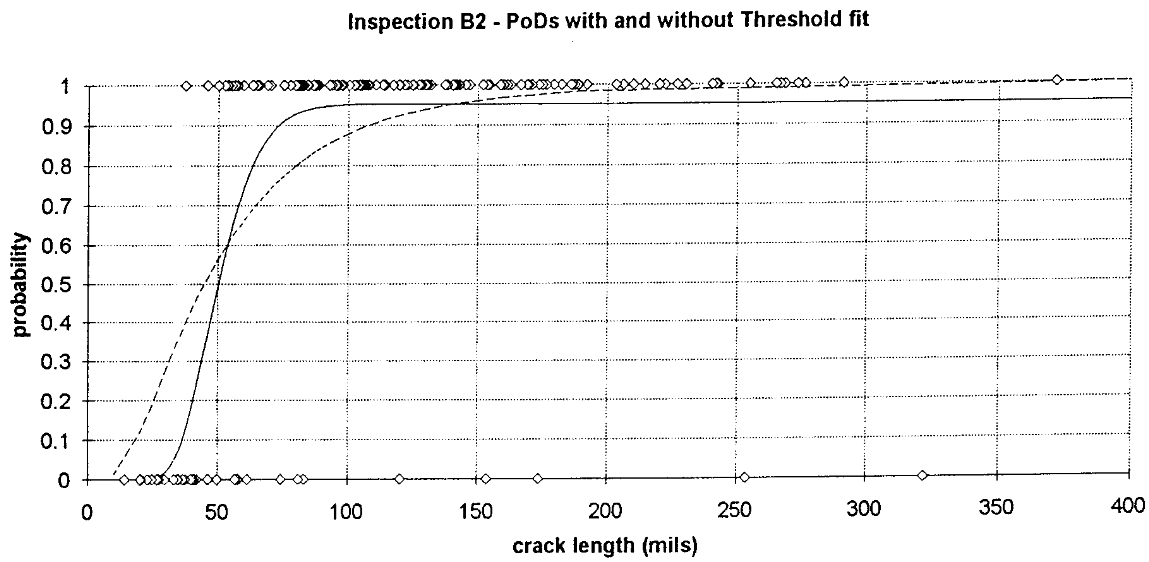

FIGURE B.2 THRESHOLD AND NON-THRESHOLD MODELS FOR INSPECTION B2.

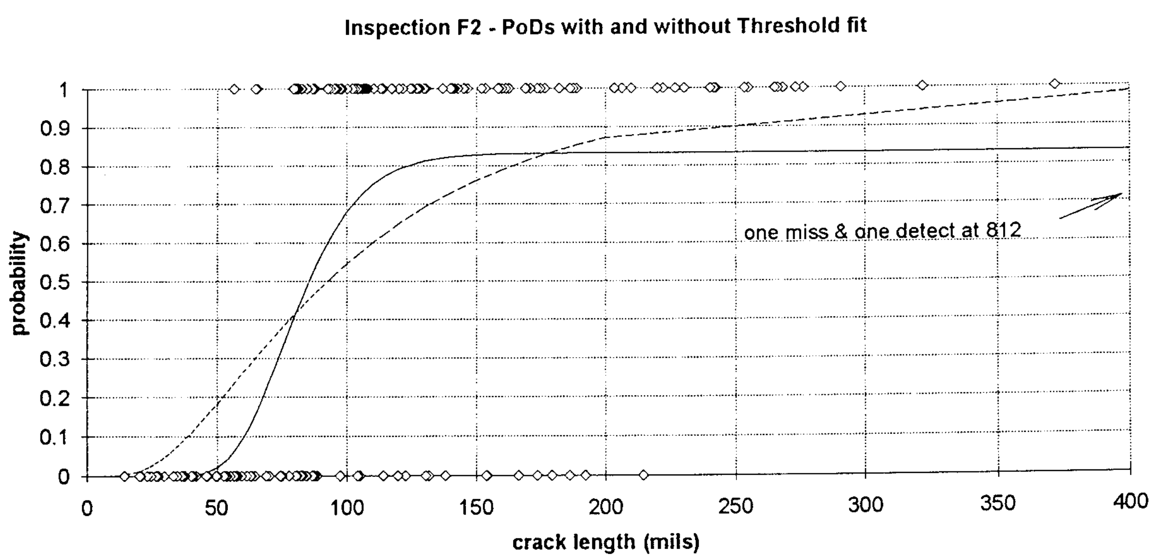

FIGURE B.3 THRESHOLD AND NON-THRESHOLD MODELS FOR INSPECTION F2.

ROC Background Data

Table B.2 gives the background on each inspector or inspection team's use of the subjective rating system in making positive calls. Also given are the number of false call for each of the subjective ratings.

Table B.2 Inspector's Use of Subjective Ratings

	total # of calls using			% of calls that are 1s	# of False Calls			% of total calls that are false			Total False
	1s	2s	3s	or 2s	1s	2s	3s	1s	2s	3s	calls
A1R	0	0	137	0.0	0	0	0			0	0
A3	0	0	144	0.0	0	0	0			0	0
B4	0	0	136	0.0	0	0	0			0	0
E2	0	0	129	0.0	0	0	3			2	3
E2R	0	0	131	0.0	0	0	6			5	6
E4	0	0	186	0.0	0	0	53			28	53
G1R	0	0	98	0.0	0	0	4			4	4
H2	0	0	129	0.0	0	0	0			0	0
H2R	0	0	138	0.0	0	0	0			0	0
H3	0	0	122	0.0	0	0	3			2	3
H1	0	1	151	0.7	0	1	16		100	11	17
A2	0	1	145	0.7	0	0	0		0	0	0
E1	2	0	132	1.5	2	0	0	100		0	2
F3	1	3	147	2.6	0	0	0	0	0	0	0
B1	3	1	146	2.7	1	0	0	33	0	0	1
C1	2	2	137	2.8	1	0	1	50	0	1	2
D2R	4	0	137	2.8	1	0	2	25		1	3
J1R	1	4	142	3.4	0	3	8	0	75	6	11
C4	3	1	103	3.7	0	0	1	0	0	1	1
A4	0	8	138	5.5	0	3	0		38	0	3
B4R	0	8	136	5.6	0	0	0		0	0	0
B3	4	4	133	5.7	1	1	0	25	25	0	2
J3	0	9	134	6.3	0	2	1		22	1	3
F4	3	7	145	6.5	0	1	2	0	14	1	3
E3	1	8	128	6.6	1	1	0	100	13	0	2
B2	7	3	142	6.6	0	1	0	0	33	0	1
J1	1	7	110	6.8	0	1	0	0	14	0	1
F1	0	12	147	7.5	0	7	2		58	1	9
C2	7	5	135	8.2	0	0	1	0	0	1	1
G2	6	5	121	8.3	3	2	1	50	40	1	6
F4R	9	6	149	9.1	8	4	1	89	67	1	13
G1	4	5	87	9.4	1	0	4	25	0	5	5
C3	11	4	144	9.4	4	2	0	36	50	0	6
D1	6	7	121	9.7	2	1	0	33	14	0	3

table continues

Table B.2 (continued)

	total # of calls using			% of calls that are 1s or 2s	# of False Calls			% of total calls that are false			Total False calls
	1s	2s	3s		1s	2s	3s	1s	2s	3s	
A1	10	8	143	11.2	9	8	5	90	100	3	22
C4R	12	2	105	11.8	0	0	0	0	0	0	0
J4	6	12	120	13.0	0	0	0	0	0	0	0
D2	12	6	119	13.1	5	1	0	42	17	0	6
J2	8	13	121	14.8	5	5	15	63	38	12	25
H4	13	10	127	15.3	9	7	8	69	70	6	24
G3	5	9	75	15.7	4	3	3	80	33	4	10
D3	15	14	145	16.7	13	10	6	87	71	4	29
F2	3	19	83	21.0	0	1	1	0	5	1	2
D4	29	73	129	44.2	21	55	8	72	75	6	84
G4	75	34	74	59.6	44	13	2	59	38	3	59

Details for the Analysis for Factor Influences

The analysis is easier to understand with a knowledge of the data structure used. A data record exists for each of the 45 inspections of the 184 known flaws. Thus, the data base consists of 8280 (45×184) records. Besides the flaw length and the rating (no detect, 1, 2, or 3) given by the inspector, each record contains information concerning the factors discussed in Section 2.2. How each factor is included in the data base is discussed below.

Test specimen type - The variable "spec" was set to 0 for the cracks on the skin panels. For cracks on the aircraft panels, spec = 1.

Crack length - Each record contained the variable "length" which was the length in mils of the longest crack present.

Off-angle condition - The variable "angle" was set to 0 , 11 or 22 according to how far off-horizontal the cracks were grown. For all the cracks on the aircraft panels, angle equals 0.

Inspection surface - The variable "paint" was set to 1 if the inspection surface was painted at the time of inspection. Otherwise, it was set to 0.

Accessibility - The variable "pos" was used to indicate the crack location at the time of inspection. The variable was set to 0 if the crack was in the upper row and it was set to 1 if the crack was in the lower row. It was set to 0 for all cracks on the large aircraft panels.

Inspection time - This factor was translated into areas of the inspection task in order that a consistency could be maintained across inspectors regardless of the total inspection time. The variable "Tseq" was defined to reflect the task sequence. This was done by breaking the skin specimens into four (4) task areas. The four task areas correspond to the upper row being divided into halves and the lower row being divided into halves. Each of the aircraft panels were considered a task area. Tseq takes on the values 1 to 6 according to the order of inspection. For example, consider the case of an inspector starting on the top row, followed by the bottom row, and culminating with the painted aircraft panel and then the unpainted aircraft panel. In this case,

for all cracks contained in the first half of the upper row, Tseq=1. For the flaws in the second half, Tseq=2. The pattern continues until Tseq=6 for those cracks on the unpainted aircraft panel.

Shift work - If the inspection occurred during the day shift then shift=1. If the inspection was done on a swing shift then shift=2. Finally, the graveyard shift was denoted by a 3.

Crack density - The variable "density" took on the value of "hi" if the crack was located on a skin panel fabricated to have more than 6 flawed sites. The value of "lo" was assigned to those cracks on skin panels with 3 or fewer flawed sites. For the cracks on the aircraft panels, density = "hi".

Three additional factors were defined. The variable "flaw_loc" took on the values of B, R, or L, according to whether the flawed location contained cracks from both sides of the rivet, right side only, or left side only. The variable "team" took on the values of 1 or 2 according to whether the inspection took place with a single person or with 2 people. The variable "proc" contained the values R, S, or T according to whether the procedure used for inspection was a rotating probe, sliding probe, or template procedure.

By including the above factors, 38 parameters were fit for model (3) in Section 4.4. There are 18 additional α parameters and 18 addition β parameters. Each of the factors adds 2k parameters, where k is one less than the number of levels contained within the factor. For example, Tseq (the task sequence or inspection time variable) takes on 6 values, therefore 5 additional α parameters and 5 additional β parameters are added to the model to allow for differences from the base case.

Initially, the data are used to estimate the full set of parameters. To estimate the potential effect of each level of a factor on the shape parameter, β, a variable was defined equal to log(crack length) when a factor was present, and equal to 0 otherwise. For example, when Tseq = 1, the variable Tseqlev1 was defined as $\log(a)$, where a is the crack length. For all other values of Tseq, Tseqlev1=0. Similar variables were defined for the other levels of Tseq as well as for the various levels of the other factors. Table B.3 gives the initial fits. The Probit procedure fits the model $C + (1\text{-}C)\Phi(\alpha + \beta log(a))$ to binary data. In order to fit a PoD curve, the procedure is run for the probability of a miss and then subtracted from 1. The resultant PoD curve is then $1\text{-}[C + (1\text{-}C)\Phi(\alpha + \beta log(a))] = (1\text{-}C)(1\text{-}\Phi(\alpha + \beta log(a))) = (1\text{-}C)\Phi({}^{-}\alpha + {}^{-}\beta log(a))$. Thus, the estimates for the α and β parameters are the negatives of the estimates for the model given in Section 4.1.1. The mean and standard deviations of the normal are $\mu = -\alpha/\beta$ and $\sigma = -1/\beta$.

Table B.3 Probit Estimates - Full Model

Log Likelihood for		NORMAL -2222.45				
Shape parameters						
Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
Log(Length)	1	-3.0904	0.7141	18.7270	0.0001	
TSEQlev1	1	-0.4027	0.4096	0.9666	0.3255	
TSEQlev2	1	-0.5447	0.3552	2.3519	0.1251	
TSEQlev3	1	-0.2602	0.3365	0.5979	0.4394	
TSEQlev4	1	0.3676	0.3163	1.3507	0.2452	
TSEQlev5	1	0.8523	0.3004	8.0482	0.0046	
SPEClev0	1	0.7893	0.3671	4.6230	0.0315	
ANGlev0	1	-0.7264	0.2522	8.2955	0.0040	
ANGlev11	1	-0.7611	0.3651	4.3465	0.0371	
PAINTlev0	1	-0.9341	0.2454	14.4852	0.0001	
SHIFTlev1	1	-0.5094	0.3153	2.6101	0.1062	
SHIFTlev2	1	-0.6659	0.3472	3.6786	0.0551	
POSlev0	1	0.3361	0.2122	2.5073	0.1133	
FLAWlevB	1	0.2853	0.3159	0.8158	0.3664	
FLAWlevL	1	-0.2604	0.2340	1.2379	0.2659	
TEAMlev1	1	0.6259	0.3335	3.5221	0.0606	
PROClevS	1	0.4482	0.3063	2.1412	0.1434	
DENSlevH	1	-0.8084	0.2375	11.5897	0.0007	

table continues

Table B.3 (continued)

Location parameters						
Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
INTERCPT	1	13.8708	3.1072	19.9270	0.0001	Intercept
TSEQ	5			25.0101	0.0001	
	1	1.8689	1.8016	1.0761	0.2996	1
	1	2.4847	1.5481	2.5762	0.1085	2
	1	1.0545	1.4413	0.5353	0.4644	3
	1	-1.5515	1.3678	1.2866	0.2567	4
	1	-3.6041	1.3182	7.4756	0.0063	5
	0	0	0	.	.	6
SPEC	1			5.0117	0.0252	
	1	-3.6066	1.6110	5.0117	0.0252	0
	0	0	0	.	.	1
ANGLE	2	.		6.8302	0.0329	
	1	2.6520	1.0828	5.9983	0.0143	0
	1	3.2572	1.5782	4.2594	0.0390	11
	0	0	0	.	.	22
PAINT	1			11.4594	0.0007	
	1	3.5672	1.0538	11.4594	0.0007	0
	0	0	0	.	.	1
SHIFT	2			4.1816	0.1236	
	1	2.3351	1.3575	2.9590	0.0854	1
	1	3.0503	1.4960	4.1576	0.0414	2
	0	0	0	.	.	3
POS	1			3.1094	0.0778	
	1	-1.6081	0.9120	3.1094	0.0778	0
	0	0	0	.	.	1
FLAW_LOC	2			4.3561	0.1133	
	1	-1.7133	1.3672	1.5703	0.2102	B
	1	1.0728	1.0032	1.1435	0.2849	L
	0	0	0	.	.	R
TEAM	1			3.2926	0.0696	
	1	-2.6076	1.4370	3.2926	0.0696	1
	0	0	0	.	.	2
PROC	2			12.1961	0.0022	
	1	-0.5256	0.1708	9.4671	0.0021	R
	1	-2.2794	1.3297	2.9386	0.0865	S
	0	0	0	.	.	T
DENSITY	1			11.9954	0.0005	
	1	3.5395	1.0220	11.9954	0.0005	hi
	0	0	0	.	.	lo
C	1	0.0214	0.0028			Lower threshold

The estimates of location parameters shifts (shifts in α) and the estimates of shape parameter changes (shifts in β) for a given explanatory variable are highly correlated. For example, the estimate for the change in β for Shift=2 is $^{-}$.6659, with a standard error estimate of .3472. The

estimate for the change in α is 3.0503, with an estimated standard error of 1.496. Both estimates are marginally significant, but they have a high negative correlation.

In order to judge whether the variable "shift" is significant, we look at the model with the Shift factor removed. It is well known that the difference in twice the Log Likelihoods of nested models [reference B1] has an approximate chi-square distribution with degrees of freedom equal to the difference in the number of parameters. Each variable is removed from the model and the resulting model is fit. Twice the decrease in log likelihoods is compared with a Chi-square distribution to judge the significance of the factor that was removed from the model. The results are given in table B.4.

Table B.4 Nested Models Comparisons- All Data

Factor	log likelihood	Chi-square	df	Pr>Chi
Original model	-2222.45			
Task sequence	-2238.07	31.24	10	0.0005
Specimen type	-2226.27	7.64	2	0.0219
Angle	-2247.98	51.04	4	0.0000
Surface (paint)	-2258.84	72.78	2	0.0000
Shift	-2225.82	6.74	4	0.1503
Position (Accessibility)	-2227.30	9.70	2	0.0078
Flaw location	-2238.86	32.80	4	0.0000
Team	-2224.55	4.20	2	0.1225
[1]Procedure	-2235.27	25.64	3	0.0000
Density	-2228.28	11.58	2	0.0031
Shift and Team	-2228.63	12.36	6	0.0544

[1]Procedure - The original model did not contain a shape factor for the rotating probe. Thus, there are only 3 degrees of freedom in excluding this variable.

Only the Shift and Team factors were not significant (at significance levels < .05). Thus, both of these factors were removed from the model and the resultant change in log likelihoods from removing both is also given in table B.4.

Since all the other factors were significant in explaining the detection data, the next step taken was to eliminate shape factors that were not significant. Several of the factors with more than 2 levels displayed significant effects at a single level. In these cases the non-significant levels were collapsed into a single level and estimates were derived. The final results are given in table B.5.

Table B.5 Parameter Estimates for Reduced Model - All Data

Shape parameters						
Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
Log(Length)	1	-3.7024	0.3623	104.4061	0.0001	
TSEQlev5	1	0.9347	0.2256	17.1589	0.0001	
SPEClev0	1	0.8479	0.3049	7.7355	0.0054	
PAINTlev0	1	-0.7592	0.2153	12.4304	0.0004	
DENSlevH	1	-0.7006	0.2305	9.2379	0.0024	

table continued

Table B.5 (continued)

Location parameters						
Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
INTERCPT	1	16.3882	1.5954	105.5232	0.0001	Intercept
TSEQ	1			16.0178	0.0001	
	1	-3.9735	0.9928	16.0178	0.0001	5
	0	0.0000	0.0000	.	.	6 (all levels)
SPEC	1			8.4603	0.0036	
	1	-3.9304	1.3513	8.4603	0.0036	0
	0	0.0000	0.0000	.	.	1
ANGLE	1			44.5368	0.0001	
	1	-0.4256	0.0638	44.5368	0.0001	hor
	0	0.0000	0.0000	.	.	off (11& 22)
PAINT	1			9.2010	0.0024	
	1	2.8001	0.9231	9.2010	0.0024	0 (bare)
	0	0.0000	0.0000	.	.	1 (paint)
POS	1			8.3741	0.0038	
	1	-0.1820	0.0629	8.3741	0.0038	0 (upper)
	0	0.0000	0.0000	.	.	1(lower)
FNUM	1			26.2610	0.0001	
	1	0.4153	0.0810	26.2610	0.0001	1
	0	0.0000	0.0000	.	.	2
PROC	2			67.8584	0.0001	
	1	-0.5346	0.1749	9.3445	0.0022	R
	1	-0.4273	0.0534	63.9027	0.0001	S
	0	0.0000	0.0000	.	.	T
DENSITY	1			9.4350	0.0021	
	1	3.0320	0.9871	9.4350	0.0021	hi
	0	0.0000	0.0000	.	.	lo
C	1	0.0230	0.0027			Lower threshold

From the model given in table B.5 it is seen that the Task Sequence level 5 was significant. The other levels of the task sequence variable were not significant, nor were there any patterns in the estimates from the other levels. The intent of looking at the Task sequence variable was to determine if there were systematic differences from the beginning of an inspection to the end of an inspection across the various inspectors. To maintain a balance in the inspection tasks, we had half of the inspectors start with the aircraft panels and the other half start with the skin specimens. Because of the way we defined the Task sequence variable, this means that half of the level 5 (Tseq=5) inspections were on the large aircraft panels. It turns out that several of the poorer performing inspectors inspected the aircraft panels last and specifically the painted panel with the large number of cracks constituted Tseq = 5. Thus the results of a few inspectors on a specific panel are influencing the results.

We consider the effect of the two types of specimens. The significance level of specimen type from table B.4 was .02. In table B.6 the parameter estimates for single curves representing all the

data are given, along with the estimates when attention is restricted to the different specimen types. The curves for all the data were shown graphically in Figure 4.26 In Figure B.4 the PoD curve fits restricted to each specimen type are shown. There is very little difference in the graphs. A formal comparison of the two curves shows no statistical difference.

It may seem like a contradiction that from one analysis (as reflected in tables B.4 and B.5), specimen is a significant factor. However, from a second analysis (as reflected in table B.6 and the curves of Figure B.4) there is no significant difference between the two type of specimens. However, the first analysis accounted for the ability of the factor specimen to explain residual variation in the presence of other explanatory factors. The parameter fits of table B.6 do not take into account any other possible explanatory factors.

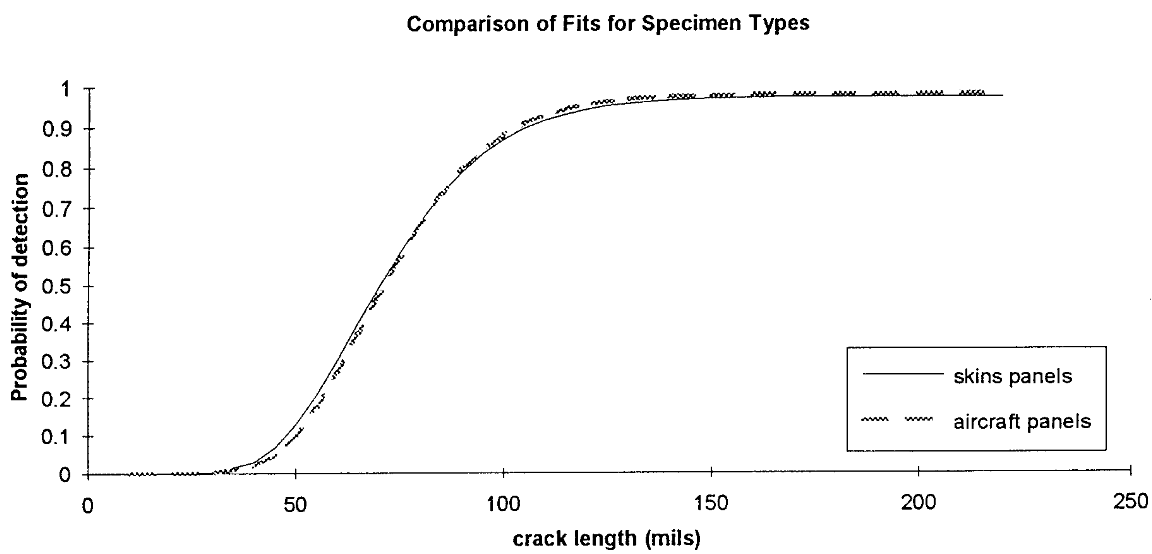

FIGURE B.4 POD CURVES FOR TWO SPECIMEN TYPES

Table B.6 Parameter Estimates Comparing Specimen Types

threshold est.	$\hat{\mu}$	$\hat{\sigma}$	var($\hat{\mu}$)	cov($\hat{\mu}$,$\hat{\sigma}$)	var($\hat{\sigma}$)	log likelihood
Complete data						
none	4.2466	.3923	7.72E-5	-3.33E-5	6.68E-5	-2520.9
.023	4.2434	.2905	5.87E-5	-1.38E-5	6.32E-5	-2384.4
			var($\hat{c}$) = 8.34E-6	cov($\hat{\mu}$,$\hat{c}$) = 5.17E-6	cov($\hat{c}$,$\hat{\sigma}$) = 7.75E-6	
skin specimens						
.025	4.2421	.2943	8.3E-5	3.13E-7	8.83E-5	-1562.39
			var($\hat{c}$) = 1.37E-5	cov($\hat{\mu}$,$\hat{c}$) = 9.32E-6	cov($\hat{c}$,$\hat{\sigma}$) = 1.03E-5	
aircraft panels						
.019	4.2601	.2732	3.39E-4	-2.83E-4	4.51E-4	-820.47
			var($\hat{c}$) = 2.57E-5	cov($\hat{\mu}$,$\hat{c}$) = -1.03E-5	cov($\hat{c}$,$\hat{\sigma}$) = 5.51E-5	

To remove the ambiguity concerning the significance of Tseq=5 we redid the analysis using the data taken from the skin specimens only. In this case, the variable Tseq has only four (4) levels and the variable spec (for specimen type) has only one value.

Table B.7 Nested Models Comparisons- Skin Specimens

Factor	log likelihood	Chi-square	df	Pr>Chi
Original model	-1457.24			
Task sequence	-1459.78	5.09	6	0.5323
Angle	-1480.48	46.48	4	0.0000
Surface (paint)	-1481.09	47.70	2	0.0000
Shift	-1461.34	8.2	4	0.0845
Position (Accessibility)	-1463.11	11.76	2	0.0028
Flaw location	-1466.25	18.02	4	0.0012
Team	-1458.10	1.74	2	0.4190
[1]Procedure	-1465.41	16.34	3	0.0010
Density	-1460.78	7.08	2	0.0290
Shift, Team and Task sequence	-1465.72	16.98	12	0.1504

From table B.7, it can be seen that the Task sequence, Shift, and Team variables are not significant. All three factors are left out and from the resultant model the shape factors that were not significant are identified and eliminated. The result is a model with the shape factors of Paintlev0 and DenslevH as significant.

As before, some levels of factors are consolidated. Interactions between the remaining factors are added to check for significance. Three interaction terms are identified as important. They are paint*procedure, paint*flaw number, and flaw number*density. Table B.8 gives the final fit which was also given in Section 4.4.

The paint (or surface condition) factor does not look to be significant. However, two interaction terms have been added into the model that involve the factor paint. Therefore, the paint factor should not be eliminated.

Table B.8 Parameter Estimates for Reduced Model -Skin Specimens

Shape parameters						
Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
Log(Length)	1	-2.7532	0.2105	171.1263	0.0001	
PAINTlev0	1	-0.6233	0.2329	7.1653	0.0074	
DENSlevH	1	-0.7049	0.2326	9.1802	0.0024	
INTERCPT	1	12.0751	.9471	162.5613	0.0001	Intercept
ANGLE	1			45.7146	0.0001	
	1	-0.4313	0.0638	45.7146	0.0001	hor
	0	0.0000	0.0000	.	.	off (11& 22)
PAINT	1			2.2944	0.1298	
	1	1.5227	1.0052	2.2944	0.1298	0 (bare)
	0	0.0000	0.0000	.	.	1 (paint)
POS	1			6.8919	0.0087	
	1	-0.1637	0.0624	6.8919	0.0087	0 (upper)
	0	0.0000	0.0000	.	.	1(lower)
FNUM	1			7.2806	0.0070	
	1	0.4546	0.1685	7.2806	0.0070	1
	0	0.0000	0.0000	.	.	2
PROC	2			63.5316	0.0001	
	1	-0.6342	0.2131	8.8564	0.0029	R
	1	-0.6622	0.0848	60.9356	0.0001	S
	0	0.0000	0.0000	.	.	T
DENSITY	1			10.6601	0.0011	
	1	3.3598	1.0290	10.6601	0.0011	hi
	0	0.0000	0.0000	.	.	lo
Paint_0*Proc_S	1			24.8052	0.0001	
	0	0.6634	0.1332	24.8052	0.0001	Surface-bare* Procedure S
	0	0.0000	0.0000			
Paint_0*Flaw #	1			4.7380	0.0295	
	1	0.4392	0.2018	4.7380	0.0295	Surface-bare *Flaw #=1
	0	0.0000	0.0000			
Flaw#*Density	1			3.4382	0.0637	
	1	-0.3660	0.1974	3.4382	0.0637	Flaw#=1 * density=hi
	0	0.0000	0.0000			
C	1	0.0244	0.0034			Lower threshold

__Retrospective Analysis with Analysis of Variance__

Table B.9 contains the individual inspection parameter fits and inspector background data. The ratings (1 to 5) correspond to those given in the protocols in Appendix B (RAE/7/Ins) or are explained in the notes to the table. The first use of this data is to estimate variance components attributable to repeat inspections, between inspectors variations, and between facilities.

The variance components are derived by modeling surface (bare or painted) as a fixed factor. Facility is then modeled as nested within surface. Thus, the facility variation will be estimated from the variation of the group of facilities doing the inspection on bare surfaces (4 facilities - 3 degrees of freedom) combined with the estimate from the group doing the inspection on a painted surfaces (5 facilities - 4 degrees of freedom). Inspector variation is estimated by nesting inspectors within facilities. There were four (4) inspectors (3 degrees of freedom) for each of the nine facilities. The error term in this model is determined from the repeat inspections. Tables B.10 and B.10a give the Analysis of Variance tables from which the variance components (given in Section 4.5) are calculated for $log(a_{50})$ and $log(a_{90})$ as estimated from the individual fits with threshold.

Table B.9 Individual Fits and Background on 45 Inspections

Insp.	$\hat{\mu}$	$\hat{\sigma}$	Background miss rate (C)	$\ln(a_{50})$	$\ln(a_{90})$	surface	shift	Team	Procedure	Freq. (kHz.)	Elapsed time (minutes)	Time on break (minutes)	Inspection Time (minutes)
A1	4.2571	0.1247	0.0097	4.2586	4.4234	B	1	1	T	200	375	88	287
A1R	4.2922	0.1337	0.0000	4.2922	4.4636	B	1	1	T	200	292	82	210
A2	4.1742	0.1000	0.0000	4.1742	4.3024	B	1	1	T	200	324	55	269
A3	4.1798	0.1199	0.0096	4.1813	4.3396	B	3	1	T	200	253	40	213
A4	4.1580	0.0565	0.0211	4.1595	4.2371	B	2	1	T	200	293	118	175
B1	4.0828	0.2016	0.0000	4.0828	4.3411	P	2	2	S	30	206	70	136
B2	3.9016	0.2454	0.0476	3.9170	4.2938	P	3	2	S	28	175	25	150
B3	4.2282	0.1846	0.0098	4.2305	4.4744	P	1	2	S	32	277	125	152
B4	4.2757	0.1885	0.0000	4.2757	4.5173	P	1	2	S	25	335	131	204
B4R	4.1138	0.1926	0.0267	4.1204	4.3906	P	1	2	S	30	318	173	145
C1	4.2644	0.1342	0.0000	4.2644	4.4364	B	1	2	S	38	314	95	219
C2	4.1311	0.1849	0.0092	4.1332	4.3771	B	2	2	S	38	292	85	207
C3	3.9733	0.2353	0.0150	3.9778	4.2942	B	1	2	S	38	340	100	240
C4	4.5340	0.3249	0.0000	4.5340	4.9503	B	2	2	S	38	354	109	245
C4R	4.4084	0.1401	0.0428	4.4162	4.6265	B	2	2	S	38	253	93	160
D1	4.3057	0.2512	0.0000	4.3057	4.6277	B	2	1	S	30	265	137	128
D2	4.3149	0.2031	0.0133	4.3183	4.5899	B	1	1	T	200	372	77	295
D2R	4.2274	0.2682	0.0000	4.2274	4.5711	B	1	1	T	200	320	100	220
D3	4.0736	0.3161	0.0000	4.0736	4.4786	B	3	1	S	27	425	102	323
D4	4.0623	0.3129	0.0000	4.0623	4.4633	B	1	1	T	200	428	67	361
E1	4.3521	0.1272	0.0000	4.3521	4.5151	P	2	1	T	100	194	32	162
E2	4.3078	0.2443	0.0342	4.3187	4.6718	P	1	1	T	100	177	19	158
E2R	4.3582	0.2824	0.0000	4.3582	4.7201	P	1	1	T	100	152	10	142
E3	4.2897	0.1299	0.0180	4.2927	4.4692	P	2	1	T	100	163	38	125
E4	4.2423	0.3339	0.0000	4.2423	4.6701	P	1	1	T	100	199	23	176
F1	3.9702	0.3116	0.0237	3.9797	4.4118	B	2	2	S	26	152	43	109
F2	4.3849	0.2423	0.1681	4.4469	5.3988	B	3	1	S	26	202	56	146
F3	4.0439	0.1150	0.0138	4.0459	4.2000	B	1	1	S	26	339	118	221
F4	3.9847	0.2723	0.0093	3.9879	4.3472	B	1	1	S	26	268	60	208
F4R	4.0701	0.1916	0.0000	4.0701	4.3157	B	1	1	S	26	391	108	283
G1	4.6624	0.2962	0.0099	4.6661	5.0577	P	1	1	T	210	447	99	348
G1R	4.5864	0.1788	0.0733	4.6042	4.9260	P	1	1	T	210	241	50	191
G2	4.3940	0.1493	0.0000	4.3940	4.5853	P	2	1	T	210	372	157	215
G3	4.7109	0.2808	0.0664	4.7360	5.2160	P	1	1	T	210	401	74	327
G4	4.2534	0.5672	0.0000	4.2534	4.9802	P	2	1	T	210	328	60	268
H1	4.2348	0.1757	0.0378	4.2435	4.5013	P	2	2	S	24	335	96	239
H2	4.3059	0.2065	0.0323	4.3145	4.6107	P	1	2	S	24	420	100	320
H2R	4.2141	0.2358	0.0122	4.2178	4.5319	P	1	2	S	24	327	121	206
H3	4.3731	0.1627	0.0678	4.3879	4.6688	P	2	2	S	24	324	89	235
H4	4.2557	0.4319	0.0237	4.2688	4.8680	P	1	2	S	24	249	40	209
J1	4.4666	0.1561	0.0000	4.4666	4.6666	P	1	1	T	250	273	21	252
J1R	4.2717	0.2227	0.0000	4.2717	4.5571	P	1	1	T	250	301	38	263
J2	4.3306	0.2853	0.0864	4.3645	4.9507	P	1	1	T	260	389	19	370
J3	4.2566	0.1271	0.0000	4.2566	4.4196	P	1	1	T	260	447	114	333
J4	4.1877	0.3426	0.0000	4.1877	4.6268	P	1	1	R	460	325	13	312

Table B.9 (continued)

Insp.	[1]Where trained -1	[1]Where trained - 2	[2]Age 1	[2]Age 2	[3]How Often EC -1	[3]How Often EC- 2	[4]How Often lap-joints - 1	[4]How Often lap-joints - 2	How long since lap-joint insp -1 (weeks)	How long since lap-joint insp- 2 (weeks)	[5]Similar to aircraft- 1	[5]Similar to aircraft- 2
A1	3		3		1		3		12		4	
A1R	3		3		1		3		12		4	
A2	3		4		1		1		5		4	
A3	3		3		1		2		3		4	
A4	3		2		1		5		24		4	
B1	3	3	2	3	1	1	3	5	3	3.50	4	4
B2	3	3	2	3	1	1	5	3	52	24	4	6
B3	3	3	3	3	1	1	4	2	52	8	4	4
B4	3	3	3	2	1	1	3	2	24	4	4	5
B4R	3	3	3	2	1	1	3	2	24	4		4
C1	3	3	3	4	1	1	2	1	2	2	5	5
C2	3	3	3	4	1	1	2	1	4	1	5	5
C3	3	3	4	4	2	1	1	3	2	1	4	5
C4	3	3	4	4	1	1	2	1	4	1	5	
C4R	3	3	4	4	1	1	2	1	4	1	5	4
D1	3		2		2		3		16		4	
D2	3		3		2		3		4		6	
D2R	3		3		2		3		4		5	
D3	1		3		3		4		26		4	
D4	3		4		5		4		24		6	
E1	2		3		1		5		18		5	
E2	3		3		1		1		1		6	
E2R	3		3		1		1		1		6	
E3	2		3		1		5		4		4	
E4	3		3		1		3		6		5	
F1	0	0	3	3	1	6	2	N/A	24	N/A	6	5
F2	2		3		6		1		34		6	
F3	1		2		1		1		24		5	
F4	2		3		1		3		48		4	
F4R	2		3		1		3		48		4	
G1	3		4		1		1		3		4	
G1R	3		4		1		1		3		5	
G2	3		3		4		4		36		6	
G3	3		4		1		5		8		5	
G4	2		4		2		4		32		5	
H1	3	3	4	3	1	1	2	5	4	2.5	5	
H2	1	3	4	3	1	1	2	5	4	52	6	5
H2R	1	3	4	3	1	1	2	5	4	52	4	6
H3	3	2	3	4	1	1	2	1	4	1	6	4
H4	3	3	4	3	1	1	2	5	4	6	4	5
J1	3		2		1		1		1		4	
J1R	3		2		1		1		1		5	
J2	3		4		1		1		1		4	
J3	3		2		1		1		3		6	
J4	3		3		1		1		1		4	

Table B.9 (continued)

Insp.	Monitors ratings from ATTITUDE and WORK PATTERN section of RAE/7/Ins									
	Interest	Cooperative	Relaxed	Hard Working	Motivated	Careful	Easy-going	Employable	Conscientious	Work Messy
A1	3	4	5	3	3	3	5	4	3	2
A1R	3	2	5	3	2	4	5	4	3	2
A2	4	5	3	5	4	4	5	4	4	3
A3	4	4	1	4	4	3	3	3	3	3
A4	5	5	3	3	4	2	5	3	2	2
B1	3	3	2	3	2	3	3	2	4	4
B2	4	5	4	4	5	4	4	4	4	4
B3	3	4	3	3	3	4	4	3	4	4
B4	5	5	4	4	3	3	5	4	4	4
B4R	2	4	5	4	3	4	4	4	4	4
C1	4	5	4	4	4	4	5	4	4	5
C2	4	5	3	4	4	4	5	4	4	4
C3	4	5	5	4	4	5	5		5	4
C4	4	5	3	4	4	5	4	4	2	4
C4R										
D1	4	4	2	4	4	3	3	3	4	2
D2	4	5	2	3	3	4	4		5	3
D2R	4	4	3	3	3	3	2.5	3	4	4
D3	5	5	3	4	4	5	4		5	4
D4	4	3	1	2	2	5	3	2	2	1
E1	4	5	4	4	4	5	4	3	4	4
E2	4	4	3	3	3	4	4	2	3	3
E2R	4	5	5	4	4	4	5	3	4	3
E3	4	5	5	5	4	4	5	4	5	5
E4	5	5	5	5	5	5	5	4	4	4
F1	2	3	3	3	2	2	3		3	3
F2	3	4	3	3	2	2	3		3	3
F3	4	5	4	4	4	4	4	4	4	5
F4	4	4	3	3	3	4	4		4	3
F4R	3	4	4	4	4	3	4		3	3
G1	4	4	3	4	4	4	4	2	4	2
G1R	4	4	2	4	4	5	3	3	4	4
G2	4	5	2	4	4	4	4	4	3	2
G3	3	4	2	4	4	4	3	2	4	3
G4	4	5	4	5	5	4	5	2	5	4
H1	4	4	3	4	4	4	3	3	4	4
H2	4	5	4	4	4	4	4	4	5	4
H2R	4	5	4	4	4	4	4	4	5	4
H3	4	5	4	4	4	5	5	4	4	4
H4	4	5	4	4	4	3	4	3	3	4
J1	4	5	4	4	4	4	5	3	4	4
J1R	4	5	3	4	4	4	5	3	4	4
J2	4	4	2	4	4	4	4	3	4	4
J3	4	4	4	3	3	4	4	3	3	4
J4	5	5	4	4	5	5	4	4	4	5

Notes for table B.9

1. Where trained -- 1-classroom, 2-on the job, 3- both
2. Age -- 1-<20 yrs., 2-20 to 30 yrs., 3-31 to 50 yrs., 4-over 50 yrs.
3. How often use EC -- 1-daily, 3-weekly, 5-monthly
4. How often inspect lap joints -- 2-monthly, 4-yearly, 5-seldom or never
5. Inspectors rating of inspection similarity to actual aircraft -- 1-not at all alike, 3-somewhat dissimilar, 5-mostly similar, 6-very similar

Table B.10 Variance Estimation ANOVA Table by Surface - Complete Data

log(a_{50})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Surface	1	0.15337	0.15337	Var(Error)+1.4 Var(Inspt(Fac)) + 5 Var(Fac(Surface)) + Q(Surface)
Facility(surface)	7	0.48260	0.06894	Var(Error) + 1.4 Var(Inspt(Fac)) + 5 Var(Fac(Surface))
Inspector(facility)	27	0.68131	0.02523	Var(Error) + 1.2 Var(Inspt(Fac))
Error (Repeat Inspections	9	0.05344	0.00594	Var(Error)
Corrected Total	44	1.37071		
log(a_{90})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Surface	1	0.27204	0.27204	Var(Error)+1.4 Var(Inspt(Fac)) + 5 Var(Fac(Surface)) + Q(Surface)
Facility(surface)	7	0.90267	0.12895	Var(Error) + 1.4 Var(Inspt(Fac)) + 5 Var(Fac(Surface))
Inspector(facility)	27	1.74004	0.06445	Var(Error) + 1.2 Var(Inspt(Fac))
Error - Repeat Inspections	9	0.08087	0.00899	Var(Error)
Corrected Total	44	2.99562		

Table B.10a Variance Estimation ANOVA Table by Surface - F2 & Facility G removed

log(a_{50})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Surface	1	0.05668	0.05668	Var(Error) + 1.4108 Var(Inspt(Fac)) + 4.892 Var(Fac(Surface)) + Q(Surface)
Facility(surface)	6	0.26477	0.04413	Var(Error) + 1.4132 Var(Inspt(Fac)) + 4.8684 Var(Fac(Surface))
Inspector(facility)	23	0.37658	0.01637	Var(Error) + 1.2043 Var(Inspt(Fac))
Error - Repeat Inspections	8	0.05152	0.00644	Var(Error)
Corrected Total	38	0.74955		
log(a_{90})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Surface	1	0.16014	0.16014	Var(Error) + 1.4108 Var(Inspt(Fac)) + 4.892 Var(Fac(Surface)) + Q(Surface)
Facility(surface)	6	0.39488	0.06581	Var(Error) + 1.4132 Var(Inspt(Fac)) + 4.8684 Var(Fac(Surface))
Inspector(facility)	23	0.59858	0.02603	Var(Error) + 1.2043 Var(Inspt(Fac))
Error - Repeat Inspections	8	0.07220	0.00902	Var(Error)
Corrected Total	38	1.22580		

The standard deviations given in table 4.4 were calculated from the variance components estimated by setting each of the Type I Mean Squares in Tabes B.10 and B.10a equal to its expectation and solving the resultant equations.

Tables B.11 and B.11a show the resultant statistics when the facilities are grouped by procedures rather than by the surface condition of the skin specimens. Within facilities D and J several procedures were followed. Facility J was classified as using the Template procedure and facility D was classed as mixed. Tables B.11 and B.11a show the procedure effect with two degrees of freedom. The three classifications; sliding probe, template, and mixed account for the degrees of freedom.

Table B.11 Variance Estimation Anova Table by Procedure - Complete Data

log(a_{50})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Procedure	2	0.23504	0.11752	Var(Error)+1.4 Var(Inspt(Fac)) + 5 Var(Fac(Procedure)) + Q(Procedure)
Facility(procedure)	6	0.40093	0.06682	Var(Error) + 1.4 Var(Inspt(Fac)) + 5 Var(Fac(Procedure))
Inspector(facility)	27	0.68131	0.02523	Var(Error) + 1.2 Var(Inspt(Fac))
Error - Repeat Inspections	9	0.05344	0.00594	Var(Error)
Corrected Total	44	1.37071		
log(a_{90})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Procedure	2	0.13205	0.06602	Var(Error)+1.4 Var(Inspt(Fac)) + 5 Var(Fac(Procedure)) + Q(Procedure)
Facility(procedure)	6	1.04266	0.17378	Var(Error) + 1.4 Var(Inspt(Fac)) + 5 Var(Fac(Procedure))
Inspector(facility)	27	1.74004	0.06445	Var(Error) + 1.2 Var(Inspt(Fac))
Error - Repeat Inspections)	9	0.08087	0.00899	Var(Error)
Corrected Total	44	2.99562		

Table B.11a Variance Estimation Anova Table by Procedure - F2 & Facility G Removed

log(a_{50})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Procedure	2	0.08041	0.04020	Var(Error) + 1.4054 Var(Inspt(Fac)) + 4.946 Var(Fac(Procedure)) + Q(Procedure)
Facility(procedure)	5	0.24105	0.04821	Var(Error) + 1.4158 Var(Inspt(Fac)) + 4.8421 Var(Fac(Procedure))
Inspector(facility)	23	0.37658	0.01637	Var(Error) + 1.2043 Var(Inspt(Fac))
Error - Repeat Inspections	8	0.05152	0.00644	Var(Error)
Corrected Total	38	0.74955		
log(a_{90})				
source	df	Type I Sum of Squares	Type I Mean Square	Expected Mean Square
Procedure	2	0.03123	0.01562	Var(Error) + 1.4054 Var(Inspt(Fac)) + 4.946 Var(Fac(Procedure)) + Q(Procedure)
Facility(procedure)	5	0.52379	0.10476	Var(Error) + 1.4158 Var(Inspt(Fac)) + 4.8421 Var(Fac(Procedure))
Inspector(facility)	23	0.59858	0.02603	Var(Error) + 1.2043 Var(Inspt(Fac))
Error - Repeat Inspections)	8	0.07220	0.00902	Var(Error)
Corrected Total	38	1.22580		

From table B.10 the repeat inspections were used to estimate a repeatability variance component for two percentiles taken from the estimated PoD curves. We can also formally test whether the two inspections can be considered as having the same PoD curves. Comparisons are made using the no threshold fits and the method presented in Appendix D of reference [10]. The parameter estimates and the T^2 statistics are given in table B.12.

Table B.12 Parameter Estimates and Covariance Terms by Inspection

			covariance terms			
	mu	sigma	mu	mu*sigma	sigma	T^2 stat
A1	4.2402	0.2275	0.00202	-0.00072	0.00126	4.4
A1 rpt	4.2922	0.1337	0.00100	-0.00035	0.00074	
B4	4.2757	0.1885	0.00136	-0.00045	0.00110	10.8
B4 rpt	4.0967	0.3981	0.00471	-0.00213	0.00320	
C4	4.5340	0.3249	0.00178	-0.00042	0.00218	4.5
C4 rpt	4.4326	0.2790	0.00174	-0.00060	0.00164	
D2	4.3120	0.2735	0.00209	-0.00084	0.00174	2.2
D2 rpt	4.2274	0.2682	0.00230	-0.00091	0.00176	
E2	4.3266	0.3750	0.00291	-0.00114	0.00261	2.0
E2 rpt	4.3582	0.2824	0.00186	-0.00061	0.00170	
F4	3.9541	0.4043	0.00574	-0.00249	0.00356	9.2
F4 rpt	4.0701	0.1916	0.00210	-0.00065	0.00141	
G1	4.6708	0.3082	0.00153	-0.00009	0.00188	0.5
G1 rpt	4.6529	0.3522	0.00191	-0.00027	0.00222	
H2	4.3201	0.3121	0.00240	-0.00100	0.00211	8.1
H2 rpt	4.1089	0.5348	0.00645	-0.00301	0.00496	
J1	4.4666	0.1561	0.00073	-0.00016	0.00083	15.3
J1 rpt	4.2717	0.2227	0.00176	-0.00072	0.00144	

Retrospective Analysis of Section 4.5

Tables B.13, B.14, and B.15 show the Analysis of Variance tables and Wilks' Lambda statistics from the multivariate tests on $log(a_{50})$ and $log(a_{90})$ estimates, respectively. The factors EC_EXP and LSJ_REC are the factors for eddy current experience (6 categories) and recency of last lap splice inspection (months). The amount of time each inspector spent on break (TOB) is also shown.

Table B.13 Analysis of Variance Table, $\log(a_{50})$

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Surface	1	0.15281	0.15281	7.29	0.0147
Shift	2	0.01734	0.00867	0.41	0.6675
Team	1	0.08269	0.08269	3.94	0.0625
Proc	2	0.03574	0.01787	0.85	0.4431
Age	4	0.15247	0.03812	1.82	0.1695
EC_EXP	5	0.18911	0.03782	1.80	0.1630
LSJ_REC	1	0.09787	0.09787	4.67	0.0445
Tob	1	0.05192	0.05192	2.48	0.1330
Model	17	0.77995	0.04588	2.19	0.0543
Error	18	0.37753	0.02097		
Corrected Total	35	1.15748			

Table B.14 Analysis of Variance Table, log(a_{90})

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Surface	1	0.23577	0.23577	6.39	0.0211
Shift	2	0.04994	0.02497	0.68	0.5208
Team	1	0.09636	0.09636	2.61	0.1235
Proc	2	0.03834	0.01917	0.52	0.6035
Age	4	0.86038	0.21509	5.83	0.0034
EC_EXP	5	0.77066	0.15413	4.18	0.0108
LSJ_REC	1	0.01449	0.01449	0.39	0.5387
ToB	1	0.00142	0.001419	0.04	0.8468
Model	17	2.06736	0.12161	3.30	0.0080
Error	18	0.66428	0.03690		
Corrected Total	35	2.73164			

Table B.15 Wilks' Lambda Multivariate Statistics (log(a_{50}) *and*log(a_{90}))

Source	Value	F	Num DF	Den DF	Pr > F
Surface	0.6200	5.2101	2	17	0.0172
Shift	0.9078	0.4211	4	34	0.7922
Team	0.8779	1.1824	2	17	0.3305
Proc	0.8088	0.9516	4	34	0.4464
Age	0.4262	2.2600	8	34	0.0468
EC_EXP	0.2945	2.8655	10	34	0.0106
LSJ_REC	0.6836	3.9351	2	17	0.0394
ToB	0.5854	6.0205	2	17	0.0105

From tables B.13 through B.15, the factors shift, team, and process were removed. In the subsequent analysis, ToB was no longer significant and it was removed. The resultant analysis of variance tables and multivariate statistics are given in tables B.16 through B.18.

Table B.16 Analysis of Variance Table for log(a_{50}) *- Reduced Factors*

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Surface	1	0.15281	0.15281	7.26	0.0127
Age	4	0.24234	0.06059	2.88	0.0444
EC_EXP	5	0.15932	0.03186	1.51	0.2229
LSJ_REC	1	0.09766	0.09766	4.64	0.0415
Model	11	0.65214	0.05929	2.82	0.0164
Error	24	0.50535	0.02106		
Corrected Total	35	1.15748			

Table B.17 Analysis of Variance Table, log(a_{90}) *- Reduced Factors*

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Surface	1	0.23577	0.23577	8.28	0.0083
Age	4	0.65173	0.16293	5.72	0.0022
EC_EXP	5	1.15355	0.22575	8.10	0.0001
LSJ_REC	1	0.00700	0.00700	0.25	0.6247
Model	11	2.04804	0.18619	6.54	0.0001
Error	24	0.68360	0.02848		
Corrected Total	35	2.73164			

Table B.18 Wilks' Lambda Multivariate Statistics- Reduced Model

Source	Value	F	Num DF	Den DF	Pr > F
Surface	0.5583	9.0971	2	23	0.0012
Age	0.3942	3.408	8	46	0.0037
EC_EXP	0.2666	4.3095	10	46	0.0003
LSJ_REC	0.7585	3.6609	2	23	0.0417

References

B.1 Christensen, Ronald, Log-Linear Models, Springer-Verlag, 1990

B.2 McCullagh, P. and Nelder, J. A., Generalized Linear Models, 2nd Ed., Chapman and Hall, 1989

B.3 Hosmer, D. W. and Lemeshow, S., Applied Logistic Regression, John Wiley & Sons, 1989.

☆U.S. GOVERNMENT PRINTING OFFICE: 1995 - 604-081/20038